LUND UNIVERSITY

DEPARTMENT OF LINGUISTICS

General Linguistics

Phonetics

WORKING PAPERS

29.1986

LUND UNIVERSITY

DEPARTMENT OF LINGUISTICS

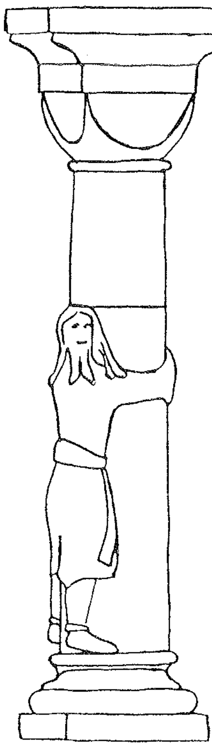
General Linguistics

Phonetics

WORKING PAPERS

29.1986

ISSN-0208-526X



Department of Linguistics and Phonetics
Helgonabacken 12
S-223 62 LUND

This issue has been edited by Eva Gårding, David House and Janna Geggus p

ANNUAL REPORT

This number of Working Papers covers a large span of topics due to contributions from both sections of our department. The varied typography is due to the different word processors and printers used.

The report will start with a list of the seminar programs.
Friday seminars (Phonetics)

Sept 13	Problems in intonation analysis: The phrase Eva Gårding
Sept 20	Report from the summer school in psycholinguistics in Brussels Ann-Christine Bredvad-Jensen and Christina Dravins
Sept 27	Experiments with a tonal grid in fluent speech Dieter Huber
Oct 4	Tone and intonation in Hausa Taghrid Anbar (Cairo)
Oct 4	Presentation of some programming problems to representatives of the department of computer engineering Mats Eeg-Olofsson
Oct 18	Acoustic-articulatory relationships in Arabic vowel production Kjell Norlin
Oct 25	Direct and indirect perception: Intonation and memory David House
Nov 8	Unorthodox tonogenesis in Hu Jan-Olof Svantesson
Nov 15	French intonation research Georges Boulakia (Paris)
Nov 22	Intrinsic pitch in Chinese Shi Bo
Dec 13	Learning to write early Caroline Liberg (Uppsala)
Jan 17	Prosodic dialectology in Fenno-Scandia Kalevi Wiik (Turku)

In January through April the Friday seminar hours were occupied by courses for advanced students. Lennart Nord's course had started in December:

Reading spectrograms
Lennart Nord (KTH Stockholm)
Intonation across several languages
Eva Gårding
Some observations on the biology of language
Alvin Liberman

Apr 14	Intonation in Finnish Kjell Weimer
Apr 18	Summary of the course "Intonation across several languages" Eva Gårding
Apr 25	The biology of language Alvin Liberman
Apr 25	Speech and the hemispheres, report on a project David Ingvar, Christina Dravins

May 2 Problems of perceptual testing
Alvin Liberman

May 30 Report from the International Conference on Acoustics, Speech and Signal Processing (ICASSP) in Tokyo
Dieter Huber

Two courses open to advanced students had been given in the fall:

Models and methods
Bengt Sigurd

A computer program for statistics: Theory and practice
Mats Nyström

The general seminar "Language, Speech, Sound & Hearing" aims at a wide audience and is conducted by the whole department in cooperation with the Department of Logopedics and Phoniatrics. It covered the following subjects.

Sept 16 Clinical voice research
Peter Kitzing (Malmö)

Sept 30 Causes behind reading and writing difficulties, report on a project
Kerstin Naucler, Eva Magnusson

Oct 7 Pragmatics and the pragmatic development of children
Carol Prutting (UCLA)

Oct 22 Expressions of politeness in Swedish
Claes-Christian Elert (Umeå)

Nov 4 Speech, language and computers: A presentation of a computer linguistics course in Gothenburg
Dieter Huber

Nov 18 The acoustics and physiology of bird song
Lars Gårding

Dec 2 The history of phonetics in Sweden
Bertil Malmberg

Dec 16 Laryngeal articulation; kinematics, control and coordination
Anders Löfqvist

Jan 17 The phonology of prosody
Jörgen Rischel (Copenhagen)

Feb 2 Situation semantics and the psychology of perception
Robin Cooper

Feb 24 Specializations for speech and other biologically significant sounds
Alvin Liberman (Yale University, University of Connecticut, Haskins laboratories)

March 3 Phonology and the problems of learning to read and write
Isabelle Liberman (University of Connecticut, Haskins laboratories

March 10 Some observations on the biology of language
Alvin Liberman

April 11-12
Humanistdagarna. Open house at the departments of the Philosophical faculty.
Theme: Emotion
Lectures:

Språk, människor och känslor (Bertil Malmberg)
Mapuche, ett minoritetsspråk i Chile
(Emilio Rivano)
Dans som språk (Jana Geggus)
Känslospråket människa-hund (Gisela Håkansson)
Intonation och känslor (Eva Gårding)
Kommunikationssvårigheter för hörselskadade
(David House)
De eller dom, en känslosak (Thore Pettersson)
Artighet, ritual i stället för känslor
(Jean-Jacques Bertout)

May 26 Unusual passives and related matters
Edward Keenan (UCLA)

The department benefited greatly from having two Fulbright scholars staying with us for the larger part of the spring term, Professors Alvin and Isabelle Liberman. Alvin Liberman gave a series of lectures on the biology of language and Isabelle Liberman lectured on dyslexia problems to the special seminar devoted to such issues. This seminar is chaired by Kerstin Nauclér and Eva Magnusson.

We also enjoyed having two Swedish Institute scholars at the department. Professor Taghrid Anbar from the University of Cairo worked with Kjell Norlin on contrastive studies of Arabic and Swedish during some weeks in the fall and Shi Bo, MA from the Acoustics Institute of Academia Sinica, Beijing, joined the program for graduate students during the whole academic year. She and Paul Touati assisted me in the course "Intonation across several languages".

There is a final report on the project "Phonetic descriptions of some important non-European languages"(HSFR) in this issue. One of the announced coming dissertations,"Language and the hemispheres", has also been sponsored by the Research Council (HSFR). The project "Reading and spelling difficulties" (HSFR) is entering its third year and "From text to prosody" (RJ, i.e. the Bank of Sweden Tercentenary Foundation) its second.

We are grateful to the Research Council for making it possible for us to buy a new graphic terminal. It is a special pleasure to acknowlege the regular support from Einar Hansen's foundation which has the aim of promoting collaboration between the universities of Lund and Copenhagen. One of the results of this exchange was a visit in January to the Dept of Audiologopedics and the Phonetics Institute at Copenhagen University by the logopedics students and their teachers David House and Sidney Wood. As usual they were extremely well received.

By the end of 1985, 26 logopedics students had completed their two-term phonetics training. As so many times before, a large group of students from the Teachers' College in Malmö were registered at our department and fulfilled the requirements for the basic course in phonetics. In addition we have had a number of students taking courses of phonetics at different levels. A group of students from Linköping University specializing in communication joined us for several days in order to get some training in experimental phonetics. A special program was designed for them.

Kurt Johansson took over the chairmanship after me on July 1, 1985 and was elected for a new three-year period in June.

Gösta Bruce spent the whole academic year in Stockholm where he substituted for Björn Lindblom. On July 1st, 1986, he succeeded me as professor of phonetics at Lund University. I wish him success in his new office.

Lund in September 1986

Eva Gårding
professor em.

CONTENTS

From prominent syllables to a skeleton of meaning: a model of prosodically guided speech recognition

Robert Bannert

ABSTRACT

In this paper an attempt is made to contribute to a better understanding of the processes involved in speech recognition. The results of some experiments constitute the basis for an outline of a model of speech recognition where prosody, especially the accentuated syllable, plays a guiding role.

In the experiments, Swedish utterances were used that were spoken with a foreign accent and that, above all, deviated tonally and rhythmically in a clear way. By introducing appropriate corrections step-by-step in a controlled manner (certain selected prosodic features in the speech signal representing the utterances were altered by LPC-synthesis), it was possible to observe the effect of these features on perception. It turned out that the word accent plays a predominant role. Accentuated syllables act as islands of clarity and stability in the speech chain due to their joint marking by a tonal change, clear spectrum, duration, and intensity. Therefore the listener is able to process these prominent and conspicuous syllables quickly and accurately. Other linguistic features (spectral, morphological, syntactic, and semantic) which are contained in the phonetically blurred and incomplete parts of the signal are subordinated to the linguistic strucutre of the identified accent pattern.

The model of prosodically guided speech recognition presented here shows some specific features. For instance, the processing units are not considered to be words in the orthographical sense (strings of letters separated by spaces on paper) but rather the accentuated syllables as anchors or fixation points surrounded by unstressed syllables. The speech recognition process does not lead to the identified meaning of a given utterance by matching a word completely analysed and with exactly defined boundaries by the acoustic-phonetic analysis with a corresponding template stored in the lexicon. The identification of meaning is instead achieved by a continuous interactive searching which is performed as an ongoing interplay between different levels of the speech recongition process. During this interplay, the current structure may be altered at any time as a consequence of several factors: new acoustic information which is continuously extracted from the incoming speech signal, linguistic constraints applicable to the intermediary structures, and the general and pragmatic knowledge of the listener.

INTRODUCTION

Perception plays a very important part in speech communication. Perception in speech can be defined as the listener's active processing of the speech signal in order to reconstruct the message intended by the speaker.

There exist a number of models of word and speech recognition. They represent different approaches and views and partly are designed in basically different ways. Often word recognition is synonymous with pattern recognition implying that the structure arrived at by the acoustic-phonetic analysis of the incoming signal is compared and matched with templates that have been stored previously. As the next step, lexical access is executed where the semantic representation, i.e. the meaning is identified. Pisoni (1984) points out particularly that a clear distinction has to be made between word recognition and lexical access. Pisoni also presents a good overview and a review of seven recent models of word recognition, among which the Phonetic Refinement Theory developed by him and his collaborators (Pisoni et al. 1985) is to be found.

The Phonetic Refinement Theory and the model of Shipman and Zue (1982) represent a real development of earlier models because they take into due consideration phonetic features and interrelationships. However, it seems to me that these models can be further developed and supplemented, especially where the role of prosody, i.e. the tonal and rhythmic aspects of speech, in word and speech recognition is conserned.

Every linguistic unit, like syllable, stress group, phrase, sentence, and text, has a specific structure, the knowledge of which is of central significance for speech recognition. The competence of the speaker/listener also contains, among other things, the knowledge of the phonotactic structure of syllables and words, their morphological structure (root, affixes), their prosodic structure, and the number of reductions and assimilations. The prosodic features are very often strongly interrelated with other phonological and morphological features, for instance phonotactic, morpho-phonological, and syntactic ones.

Models of speech perception have to cope with the fact that the speech signal is not always distinct and complete. Instead, most often the acoustic signal arriving at the listener's ear contains distortions of different kinds. These deviations appear as the consequences of at least three dimensions of indistinctness, namely of speech tempo (slow - fast), of articulation (distinct - lax), and of the linguistic distance between a norm or standard and the actual form (small - large) which contains regional, social, and individual features and foreign accent as well. These deviancies include reduced or missing spectra compared to the intended form both of which can originate from lax or fast speech of from external distortions. Opposite to imcompleteness, the signal may contain a larger number of segments, e.g. produced by vowel epenthesis. Some spectra may be analysed only to a certain extent, e.g. an [m] only as nasal, an [ø] only as palatal, prosodic features may be wrong (short instead of long vowel), the accent may be placed on the incorrect syllable (telephóne), etc. Therefore it has to be assumed that the result of the acoustic-phonetic analysis not always amounts to a complete and unambiguous phonological form which will lead directly to the lexical element which, eventually, will be identified correctly. On the contrary, the phonological representation as the result of the working of the bottom-up processes has to be thought of as incomplete and deviant compared to the meaning intended by the speaker.

An adequate model of speech perception should be able to handle a rather wide variation in the speech signal. Listeners do communicate with each other in spite of individual, social, geographical, and other differences in their pronunciation. A clear instance of large acoustic deviances is to be found in foreign accent. Thus phonetic variation is a rather complex phenomenon, resulting from a given language being spoken not only as the first language but also as the second language by people with different first languages. A simplified multi-dimensional model of phonetic variation is presented in Bannert (1982).

It is the aim of this investigation to present an outline of a model of speech recognition in which prosody plays a significant and leading role (*). This model is compatible, to certain parts, with other models of word or speech recognition, especially with the Phonetic Refinement Theory presented by Pisoni et al. (1984). However, it adds some new aspects focussing on the incompleteness and fuzziness of the results of the acoustic-phonetic analysis. It does not work with the processing unit of the word characterized by clearly defined boundaries. The phonological form of the word is being built and assembled starting from phonological fragments and applying, among other things, the morpho-phonological knowledge of the listener.

METHOD AND MODEL

As the basis for testing intelligibility of Swedish spoken with a foreign accent, the so-called correction method is used (Bannert 1978). The starting point is an utterance spoken by a non-Swedish speaker. It contains certain features which are analysed and well-defined. These deviations are corrected in the acoustic signal step-by-step by means of LPC-synthesis. The tonal feature of accent corresponds to certain parts of the Fo-contour of the utterance, the temporal features of quantity and phrase rhythm are manifested in the durations of the segments and their relationships to other segments and groups of segments (syllables). The Lund model for Swedish intonation (Bruce 1977, Gårding and Bruce 1981) served as the model for the tonal corrections. The temporal corrections had to be carried out according to estimated values which were then checked auditorily. We are still lacking a comprehensive model for Swedish speech rhythm.

Intelligibility of foreign accent is investigated against the background of a kind of Active Direct Access Model for word recognition similar to the model developed by Marslen-Wilson and Welsh (1978). Recognizing a word is considered a time-dependent active process where acoustic (bottom-up) and linguistic, pragmatic, and general information (top-down) conspire.

Intelligibility is seen as an aspect of the bottom-up processes. Intelligibility is high if the acoustic, auditory, and phonetic analysis of the speech signal results in a possible phonological structure processed in the short-term memory that easily and quickly can find its way to the phonological representation of a word stored in the long-term memory. In the opposite case, intelligiblity is low if the speech signal is analysed in such a way that no corresponding lexical element can be discovered. Thus intelligibility facilitates decoding by decreasing the demands on the top-down component and makes comprehension faster, easier and better.

If the model of Active Direct Access is expanded to foreign-accent speech, one prediction, then, will be that a longer stretch of acoustic information - a larger chunk of the speech signal - is needed before a word spoken with a foreign accent can be recognized. Thus, due to the acoustic deviations, it should also take more time to process foreign accent. Foreign accent puts a lot of strain on the short-term memory and a heavier demand on the top-down processes.

TEST PARADIGM

Testing intelligibility is not an easy task. After considering different problems, their approaches and possible solutions, a test paradigm was constructed which is shown in Fig. 1. Samples of foreign accent that are clearly deviating in the prosodic features to be investigated constitute the starting point. Using LPC-speech synthesis, each utterance is altered in such a way that corrections corresponding to the deviating prosodic features are introduced into the speech signal <1>. In this way, families of utterances are created that consist of several members, namely the original utterance and several versions that differ from the foreign accent original by a certain correction or improvment. Each family of utterances is extended by adding an idiomatic version spoken by a male Stockholm speaker. Thus a dimension of variation within each family is established where the foreign accent original and the Swedish version mark the end points and the corrected versions are assumed to lie in between.

All these utterances are then distorted in different ways, using noise and increased speech tempo <2>. This test design makes it more difficult for the listener to understand speech, and, at the same time, it is easier to discern the effects of the various corrections. In both tests, the signal-to-noise ratio in the noise test and the speech tempo in the second test were chosen based on preliminary tests using naive listeners. The listeners understood the utterances with difficulty.

The Swedish listeners who were not accustomed to foreign accent participated individually in the listening tests. They heard the test utterances via loudspeakers in the perception laboratory and repeated in their own Swedish without hesitation what they could understand of the utterances played to them <3>. The listeners were urged to respond, even if they did not understand the whole utterance and to guess freely in case of uncertainty. The test utterances and each listener's responses were recorded on different channels of a REVOX tape recorder. The responses were analysed, evaluated, and compared to the intended meaning. Response time in the noise test was measured. Transmitted prosodic information analysed in the oral responses and response time are the

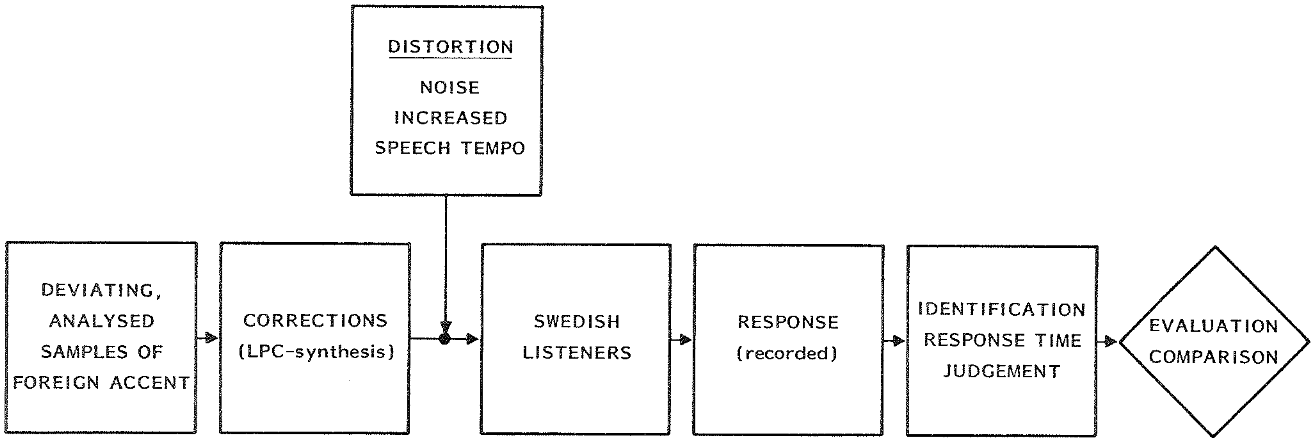


Fig. 1 The test paradigm.

basis for a ranking of the phonological features.

MATERIAL AND CORRECTIONS

The test material consisted of 9 relatively short utterances <4> the length of which ranged from a single compound word of 3 syllables (utterance 6) to a complete sentence consisting of subject, adverb, and adverbial phrase (utterance 4). The test utterances were rendered by three male speakers with French, Greek, and Persian (Farsi) as their first language (L1). These utterances and their corrections are shown in Fig. 2. The features of quantity and phrase rhythm are corrected by changing segment durations, the feature of accent <5> is corrected by inserting a tonal peak in the accentuated syllable and, at the same time, deleting the original peak in the incorrect syllable <6>. An illustration of the temporal and tonal corrections in utterance 8 is shown in Fig. 3.

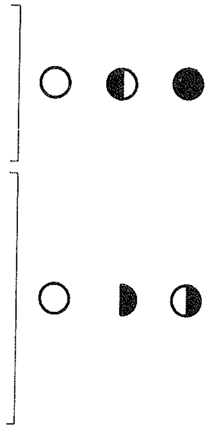
The kind of stimuli and their number varied from test to test. The test material common to all 3 tests consisted of the 9 original utterances spoken with foreign accent and three corrections each. They were interspersed with twelve different utterances spoken with foreign accent which served as distracters and, at the same time, as calibrators for the reliability of the listeners' responses. The intelligibility tests also contained a version of each of the original utterances of foreign accent spoken by a male Stockholm speaker. Furthermore, the intelligibility test presented with increased speech tempo and the acceptability test also contained a version of each of the 9 original foreign accent utterances representing deteriorated Swedish. These deteriorated Swedish utterances were produced by introducing into each of the original Swedish utterances all the prosodic deviations of its original utterances <7>. All the stimuli were re-synthesized. They were free from distortions such as clicks or buzzes and sounded quite natural.

LISTENING TESTS

Each listening test concerning intelligibility consisted of three parts. First, the utterances of the four speakers were presented in the following order: French, Greek, Persian, and

UTTERANCE No	L1	UTTERANCE
1	FRENCH	EN KAFFEBRICKA
2	"	SOM EN MYCKET LITEN POTATIS
3	"	LITE RÖDA TYGBITAR
4	GREEK	Å SOLEN LYSER BLEKT I SÖDER
5	"	BÅDA ÄR DYRA
6	PERSIAN	MARKATTA
7	"	DET ÄR EN MÅNDAGMORGON
8	"	I SAMHÄLLET

CORRECTIONS



PITCH ACCENT

PHRASE RHYTHM

QUANTITY

Fig. 2 The eight utterances and their corrections.

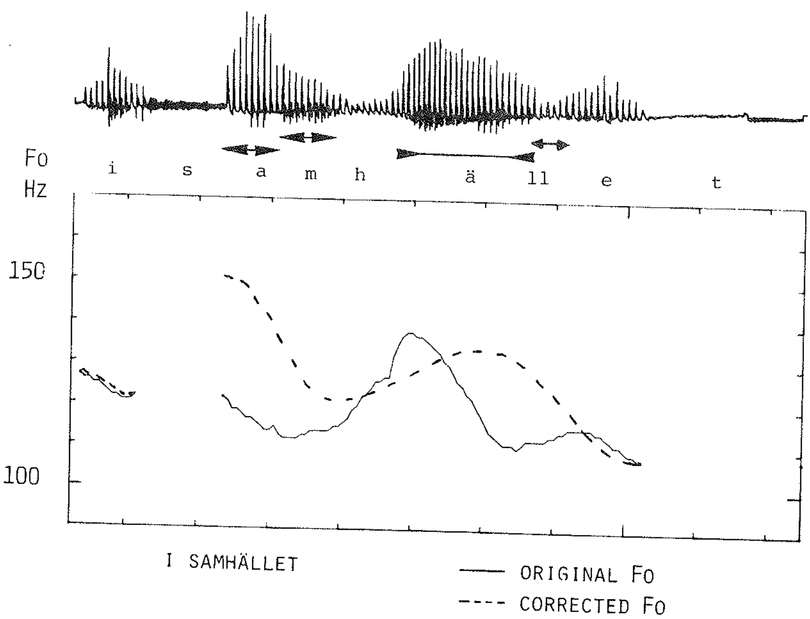


Fig. 3 An illustration of the temporal and tonal corrections.

Swedish. The speakers read aloud a text, approximately 45 seconds each. Thus the listeners were given the opportunity to adapt to the four speakers' voices and pronunciation. This speaker adaptation is also necessary as a precondition in order to be able to comprehend speech in a normal way. In the acceptability test, the Swedish speaker was exluded, as were the original Swedish utterances. Second, eight utterances (six in the acceptability test) followed, two for each speaker, presented in the same way as the test proper. The utterances of the second part were intended to get the listeners accustomed to the presentation of the utterances under noise, with increased speech tempo for the intelligibility tests or in the normal way for the acceptability test and to practice responding to the stimuli. These practice utterances did not appear in the test. Third, the noise test proper contained 21 utterances, namely nine test utterrances (1 version of each original utterance) and 12 distracters. As each listener responded to each utterance only once, the 45 stimuli [(original foreign accent utterance + 3 corrections + original Swedish version) x 9 utterances] of the whole test were divided into 5 series and administered separately to 5 different listener groups. The test with increased speech tempo contained 54 test stimuli (45 stimuli + 9 deteriorated Swedish utterances). The test was divided into 6 series and given to 6 new listener groups.

In the two intelligibility tests, each listener heard and responded to each test utterance only once. In the acceptability test three marks (1 for a low degree of foreign accent, 2 for a high degree of foreign accent, and x for a degree of foreign accent in between) were given to each stimulus by each of the 20 speakers. The total duration of the listening test under the conditions noise and increased speech tempo, respectively, was about seven minutes. The acceptability test took about 20 minutes. Fifty South Swedish listeners (university students) took the noise test individually. Thus each version was given 10 responses by the whole group. The test with increased speech tempo was taken by 30 listeners under the same conditions. Thus, in this test, each version was given 5 responses by the whole group.

The responses of the two intelligibility tests were analysed and evaluated according to a scoring system that primarily counted the prosodic information contained in the listeners' responses, such as number of accents, number of syllables (vowels), stress pattern, syllable quantity (long/short vowel and consonant), etc. The response time was defined as the time lag between the end of the test utterance and the beginning of the listener's response measured on duplexoscillogrammes. The marks of the acceptability test were counted and are given as group scores.

RESULTS

The results of the three experiments are reported as follows: The effects of various prosodic features along the dimension intelligibility are shown under the two conditions, namely original speech tempo and noise in Experiment 1 (Fig. 4) and increased speech tempo in Experiment 2 (Fig. 5). In parallel and supporting this aspect, the reaction times of Experiment 1 illustrate rather the psycholinguistic dimension of speech processing (Fig. 6). Some selected typical examples of listener responses that did not correspond to the intended utterances and which may provide some revealing information about the processes involved in speech recognition are presented and analysed. The assessment of the stimuli according to their acceptability in Experiment 3 provides some useful illustration of the psychological and subjective aspects of speech recognition (Fig. 7). Finally the results of each experiment are compared with one another: Intelligibility and Reaction time in Experiment 1, both with Intelligibility in Experiment 2, and the three of them with Acceptability in Experiment 3.

Intelligibility

In Figures 4 and 5, the distribution of the stimuli under the condition Noise (original speech tempo) and Increased speech tempo, respectively, are given as percentage along the dimension intelligibility which is defined in terms of transmitted prosodic information and represented as a straight line. Each of the 8 families of utterances <4> is shown individually.

Fig. 4 shows that the utterances with the original and uncorrected foreign accent often have only a low intelligibility which expectedly is in opposition to the Swedish corresponding utterances that are understood very well and without difficulty. However, in both cases there are several exceptions. Utterance 4 is especially conspicuous showing a relatively high degree of intelligibility in spite of its foreign accent. The corresponding Swedish utterance, on the contrary, shows a low degree of intelligibility. In

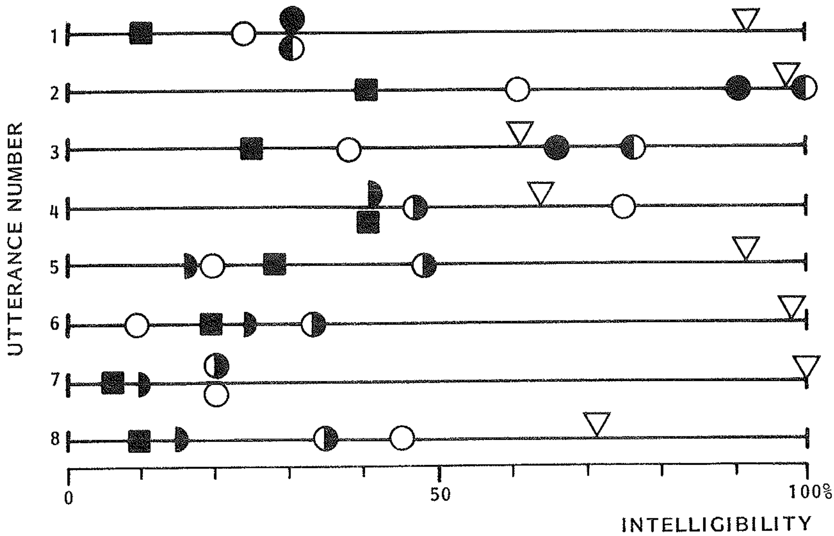


Fig. 4 The effect of the corrections on intelligibility expressed as the prosodic information contained in the listeners' responses (noise condition).

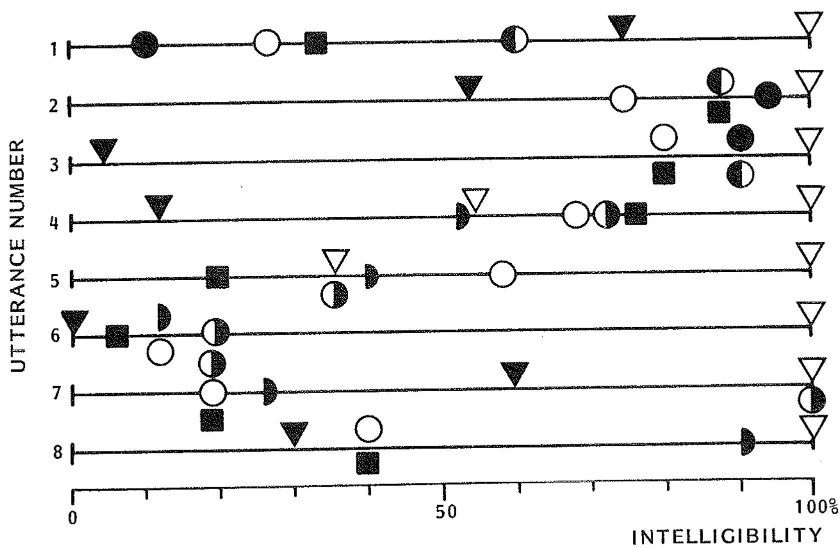


Fig. 5 The effect of the corrections on intelligibility expressed as the prosodic information contained in the listeners' responses (increased speech tempo).

most utterances, the corrections bring about an increase of the degree of intelligibility, utterances nos. 4, 5, and 6 being the exceptions. In some cases (utterances nos. 2, 3, 4), in fact, a high degree of intelligibility is reached as a consequence of the corrections. Distributed across the whole material, the corrections of word accent and (phrase-)rhyhtm produced the best results (exceptions here are utterances nos. 4 and 8).

Fig. 5 also shows the distribution of the stimuli along the dimension Intelligibility, but here the deteriorated Swedish versions are added. Compared to Fig. 4, this result, by and large, is quite similar. The utterances with the foreign accent show a relatively low degree of intelligibility, their original Swedish counterparts, on the other hand, a very high degree. Even under this condition, utterance no. 4 is the exception. The original Swedish utterances suffer from a dramatic decrease in the degree of intelligibility when the deterioration of the prosodic features are introduced into the signal. This holds especially for utterances nos. 3, 4, and 6. Under the condition of Increased speech tempo, too, the corrections increase intelligibility. And here, too, the combined correction of word accent and (phrase-)rhythm lead to the best results in most cases (clearly in utterances nos. 2, 3, 4, 6, 8).

A comparison of Figures 4 and 5 reveal minor differences. For instance, the corrections of utterances nos. 2, 3, and 8 in Fig. 5 (Experiment 2) produce higher values. These differences might be attributed above all to the different conditions in the two experiments. The behaviour of utterance no. 4 which clearly deviates from the other utterances, even with respect to reaction time (see the following section), might be attributed to its syntactic complexity and its length.

Reaction times

Fig. 6 shows the distribution of the stimuli of Experiment 1 (original speech tempo, noise) according to reaction time of the group scores. Compared to the results concerning intelligibility, by and large, similar relationships between the different versions of an utterrance are to be found. The original Swedish utterances almost always show the shortest reaction times, as could be expected. But even in this case, utterance no. 4 comes out as the real exception. In only a

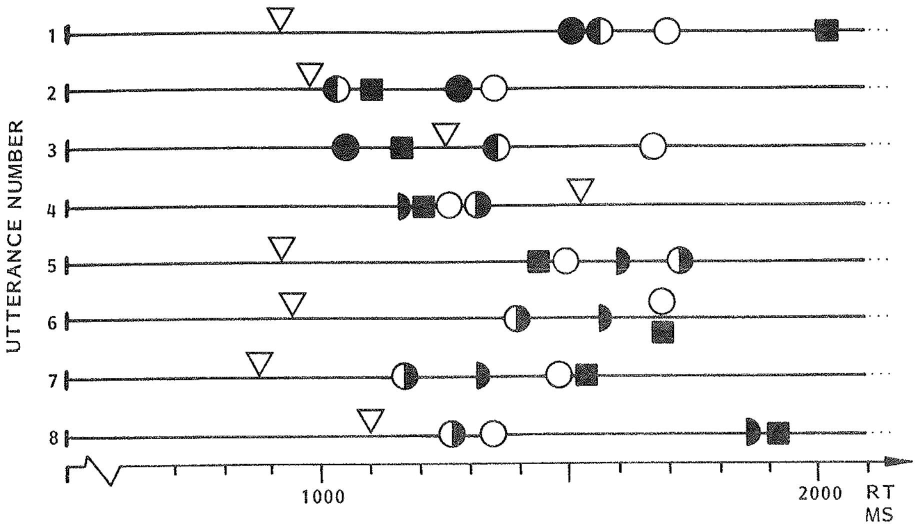


Fig. 6 Response times (noise condition).

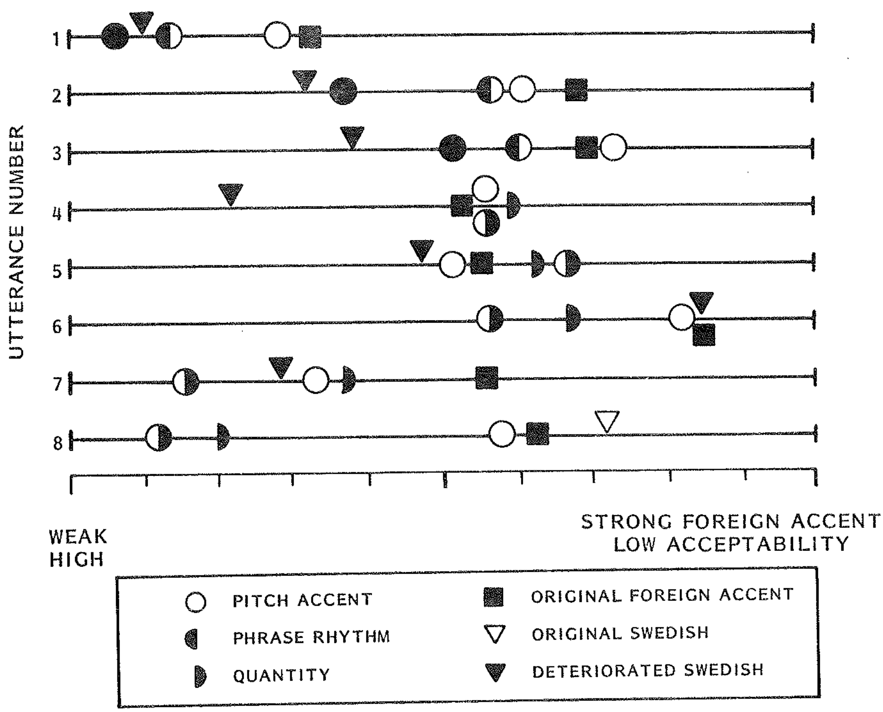


Fig. 7 Acceptability of the original foreign accent utterances, their corrections, and the deteriorated Swedish utterances.

few cases do the uncorrected foreign accent utterances have the largest reaction times. The shortest reaction times among the corrected utterances are for those stimuli in five cases where word accent and (phrase-)rhythm were corrected in combination (utterances nos. 1, 3, 6, 7, 8).

A comparison of these results concerning reaction times in relation to the corrected versions reveals that the combined correction of word accent and (phrase-)rhythm are characterized by the shortest reaction times which may be interpreted as an indication of easy and fast processing and, at the same time, bring about the highest degree of intelligibility.

Acceptability

Fig. 7 shows the distribution of the stimuli assessed by the listeners according to the degree of foreign accent which may serve as a direct measure of acceptability. The test material here is arranged along the dimension of high vs. low acceptability. In most cases, the utterances with the original foreign accent are assessed with the lowest degree of acceptability. Only utterances nos. 4 and 5 represent clear exceptions. Those corrections where the features of word accent and (phrase-)rhythm were manipulated in combination, with only one exception, namely utterance no. 5, show the highest degree of acceptability.

The deteriorated versions of the original Swedish utterances show a very low degree of acceptability in six cases (utterances nos. 1, 2, 3, 4, 5, 7), and in four cases (utterances nos. 2, 3, 4, 5), in fact, the lowest degree of all versions.

Comparing the results concerning acceptability with those concerning intelligibility under different conditions, reveals a clear and parallel behaviour of certain stimuli. Among the corrected features, the combination of word accent (accent pattern) and (phrase-)rhythm stands out as the most efficient one. These stimuli obtain the highest degree of intelligibility under different conditions, need the shortest reaction times, and are accepted most readily.

Response patterning

Analysing and evaluating listener responses with respect to the altered features of the stimulus may provide some information about the processing of the incoming signal. Those cases where the listener did not respond at all or responded with the intended utterance <8> are not very interesting.

Some typical examples of listener responses showing the effects of only correcting word accent on the results of the speech recognition processes, i.e. putting the tonal change onto the correct syllable, are given in Table 1.

In general, the accentuated syllables in the stimulus, no matter where, come through very well. The responses to the original utterrances where word accent falls on the wrong syllable correspond exactly to this accent pattern. By correcting the word accent only, i.e. by shifting the tonal change onto the right syllable, the responses change in such a way as to correspond to the new and correct accent pattern. In many cases it can be observed that the accent pattern of the incoming signal determines the accent pattern of the response which will be identical, although the other linguistic structures and features of the response differ with respect to the simulus. It seems as if spectral, morphological, syntactic, semantic, and pragmatic elements are fitted into the framework laid out by the accent pattern. Thus it appears rather clearly that the word accent syllable serves as the first and decisive sign post or guide in the processing of the acoustic information at non-peripheral levels. Therefore word-accentuated syllables, as a consequence of their prominent marking by combining tonal, rhyhtmic, spectral, and dynamic features in them, play a predominant part in speech recognition. Other linguistic aspects of the possible linguistic structure, drawing upon all kinds of information available, seem readily to be subordinated to the gross structure defined by the accent pattern.

DISCUSSION

First the effects of the corrected prosodic features on speech recognition are commented upon. Second an outline of a model of prosodically guided speech recognition will be

Table 1. Some typical listener responses

SWEDISH	FOREIGN ACCENT (ORIGINAL)	CORRECTED WORD ACCENT
en 'kaffe,bricka	(en kaffebri'cka:)	
	ja e inte 'klar	en 'vacker 'flicka
	en liten 'flicka	'kaffet e 'klart
'båda är 'dyra	(båda är 'dyra)	
	va de 'blir	'båda,dera
	på dig 'båda	'både ...
	fotogra'fi	
'mar,katta	(mar'ka:ta)	
	man 'pratar	'prata
	dom e 'korta	'marknad
	ma'kaber	
det är en 'måndag,morgon	(det är en mån'da:gmorgon)	
	det är en nou'gatmålning,	det är en 'vårmorgon,
	det är en han'garmålning	det är en 'söndagmorgon,
		det är en 'kundradio
i 'sam,hället	(i sam'hä:let)	
	i sin 'helhet	i 'sandträdet
	utan 'teve	i 'samlingen
		i 'handlingen

The accentuated syllable is marked by ' preceding it. A long vowel is specified in the very broad transcription of the original stimulus in parentheses for reasons of clarity.

presented and, third, speech recognition is discussed under the aspect of the difficulties related to foreign accent.

The effect of the corrected prosodic features on speech recognition

It was expected that the manipulations in the speech signal which were made in a controlled and step-wise way should make it possible to make clear and definite statements about the effect of the corrected prosodic features. However, the results clearly show that there is not always a simple and direct relationship between the correction of a given prosodic feature and the listeners' reaction to it. Thus it happens several times that the corrected version shows a lower degree of intelligibility than the utterance with the original foreign accent (for instance Fig. 4, utterance no. 5; Fig. 5, utterances nos. 1, 2, and 4). Corresponding statments can be made with respect to reaction times and acceptability. In the opposite case, the Swedish original utterances, too, get low scores rather often and, in fact, they score worse than some corrected versions (for instance, Fig. 4, utterances nos. 3 and 4; Fig. 5, utterance no. 4; Fig. 6, utterance no. 4).

This unexpected behaviour may have several explanations. No doubt, however, the reason can hardly be found in the fact that the score of each version in the dimensions Intelligibility, Reaction time, and Acceptability, respectively, was given by a different group of listeners. In order to eliminate this supposed factor, the number of the listeners has to be increased considerably. But, as far as the assessment of the speech signal by the listeners is concerned, it is quite clear that phonetic and phonological deviating features are analysed and evaluated rather differently. In my experience, this individual reaction pattern on behalf of the listeners always becomes obvious, even when trained and experienced teachers of Swedish as a second language are exposed to foreign accent.

However, a more plausible explanation for this divergent behaviour seems to be found on purely phonetic and phonological grounds. The various manipulations, e.g. only vowel or consonant duration, only word accent, may have a somewhat negative effect on the processing of the speech signal by the listener. This is because some manipulations may interfere with the various processes involved in speech recognition or even impair and, at worst, block them. We must not forget that even the corrected signal is clearly heard as foreign accent as it still contains certain prosodic and all of the segmental deviances. By correcting only one feature at a time, new constellations of foreign accent may be created which, against the background of the interplay between prosodic features, segmental features, morphological and syntactic features, may exert an impairing influence on the processing of the speech signal.

In conclusion, then, a general statement can be made: compared to the original Swedish utterances and also to the deteriorated Swedish utterances, most of the prosodic features affect speech recognition in a positive way by increasing the degree of intelligibility, by decreasing the reaction time, and by being accepted to a higher degree compared to their deviant counterparts. The largest positive effect is exerted by the combination of word accent and rhythm.

With respect to the significance of prosody in speech recognition, another general statement can be made: The accent pattern, rhythmic structure, and overall intonation contour facilitate purposefully the successful processing of the speech signal. These features give a macro-structure to the speech chain by dividing the spectral events or the stream of sounds into useful units larger than sounds and syllables, namely accent groups, prosodic phrases or intonation units (cf. Nespor and Vogel 1983) <9>.

Outline of the model

On the basis of the results of my investigations and the models of word recognition mentioned in the introduction, a prosodically guided model of speech recognition has been developed. Prosodic features play the decisive part for the searching of lexical elements. The model, outlined in the following in a simplified way, is shown in Fig. 8 <10>. It describes an interactive process on several levels where information and knowledge of various kinds affect the recognition process from the speech signal to the identified meaning.

Starting with the input, the acoustic-phonetic basic information of (one part of) the utterance is extracted by the peripheral auditive-acoustic analysis. This first automatic analysis proceeds from left-to-right, i.e. the incoming speech signal is processed continuously along the

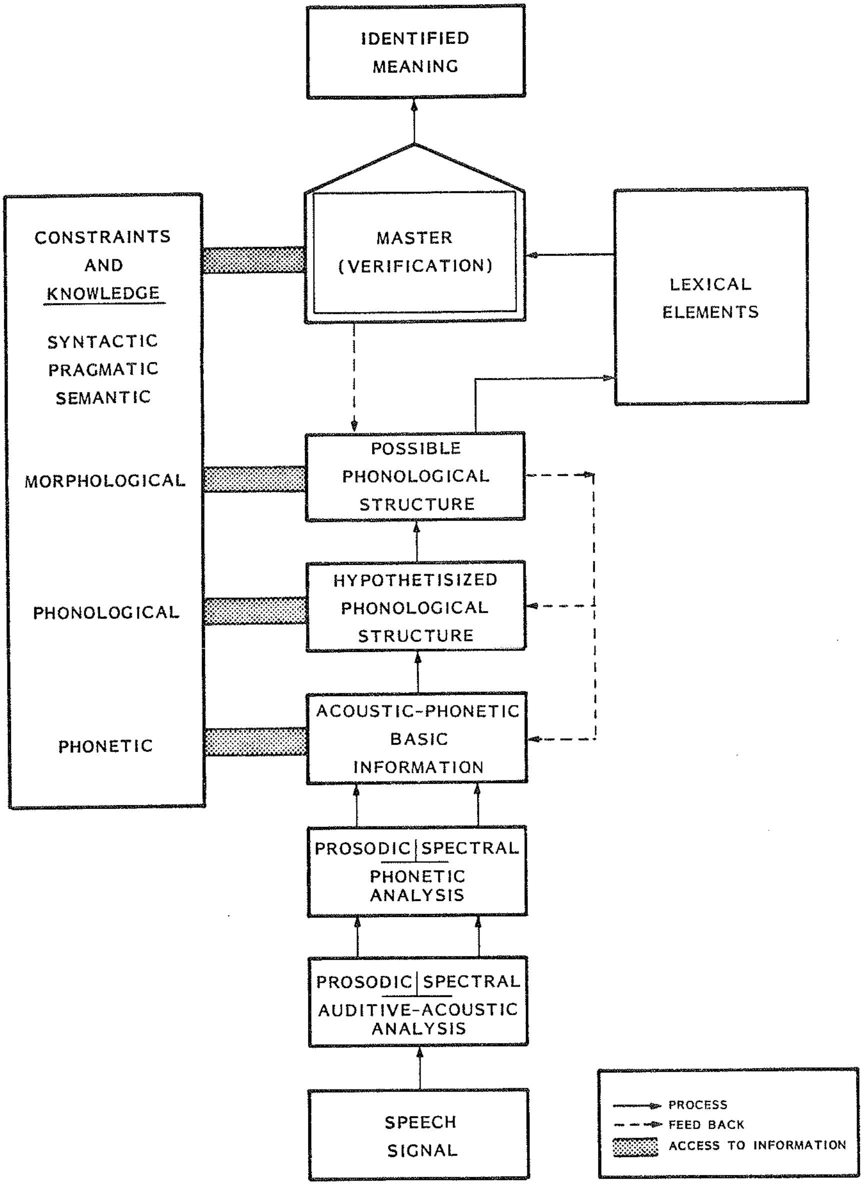


Fig. 8. A model of prosodically guided speech recognition.

time axis.

The acoustic analysis is done in two different channels, namely the prosodic and the spectral one (cf. Svensson 1974, House 1985). While in the prosodic channel the tonal and temporal features of the chunks of the processing units are established, the spectral channel provides the information about the qualitative features (formant frequencies, band widths, etc.) of the segments <11>.

Quite often the auditive-acoustic analysis cannot always result in a complete phonetic basic structure. The speech signal may be blended with acoustic distortions from outside, the signal-to-noise ratio may be too small or the speech signal might contain in some form components that are reduced, missing or, with respect to the form expected by the listener, deviant in some other way.

The auditive-acoustic analysis is followed by the phonetic analysis which combines and integrates the auditive-acoustic parameters into chunks of approximately the size of a syllable and which labels it phonetically. The phonetic labelling, most often, cannot be performed in a refined way (cf. Pisoni et al. 1984). The phonetic interpretation provides the basis for the acoustic-phonetic basic information about the chunk of the speech signal to be processed.

The acoustic-phonetic basic information is structured according to prosodic and spectral features. The prosodic features provide the position of the accentuated syllable or syllables in the chunk or chunks; the spectral features contain information about the spectral gestures of the segments. Taken together they provide information about the number of syllables in the chunks. There is, however, a clear difference between the two dimensions: while the accentuated syllable always appears correct in the basic structure, the spectral component often remains classified only in a gross manner.

This fact has certain consequences for the emergence of the hypothesized phonological basic structure on the following level: The spectral elements in the acoustic-phonetic basic information are subordinated to the prosodic structure of the accent groups where accent group means the accentuated syllable surrounded by the unstressed syllables. This subordination is brought about by the top-down constraints and the general knowledge of the listener which operate in generating the hypothesized phonological structure. These constraints are phonetic, phonological, morphological, syntactic, and semantic.

The hypothesized phonological structure is not generated only once and for ever but, instead, can be altered in a short period of time as a consequence of not only new acoustic-phonetic information but also of new top-down information which is flowing forth and thus becomes available all the time. The definite hypothesized phonological structure of accent groups generates the possible phonological strcuture of (chunks of) utterances which are stored in the Short-Term Memory (STM) as well. Now the search in the lexicon in the Long-Term Memory (LTM) for lexical elements which correspond phonologically to the equivalent elements stored in the lexicon will start. The semantic elements of the lexicon are arranged in a multi-dimensional fashion according to various phonological features and structural characteristics. These possible phonological structures provided by the analysis of the speech signal and the working of linguistic constraints, it must be assumed, normally do not look like orthographic words with clearly defined boundaries, which correspond exactly to a stored counterpart. They are not searched for like a numbered book in a bookshelf and found immediately by its distinctive digit. Approaching the lexical elements would rather amount to a search consisting of a large array of activities utilizing different features simultaneously. The possible phonological structure which emerged from the fragments of the acoustic-phonetic basic information contains the accentuated syllable as its most important search criterion which is stuffed with the most distinct acoustic and structural information. Therefore it can be assumed that the search starts out for phonological representations of lexical elements showing the identical accent pattern and most of the spectral features of the accentuated syllable. Of course, all the information concerning the surrounding syllables is used as a supporting criterion as well. In general, it has to be assumed that speech recognition is characterized by an interplay of activities where all information available is processed simultaneously and optimally. This kind of search assumes explictily that the boundaries in the possible phonological structure need not be defined exactly and in advance. The first aim of the search for lexical elements seems to be to find the syllables with the most distinct marking which, in turn, are identical with the basic meaning of the root or stem of a word, i.e. to find the skeleton or the corner stones of meaning.

As is generally known, languages use different principles for

accent distribution in their information structure. In accent languages like, for instance, Swedish, English, and German, word accent, in principle, exactly functions for signalling the word stem as the kernel of the meaning of a word. This is true both of morphologically simple and complex words. But also in languages with different principles for accent distribution, like for instance Finnish and Czech with initial accent or Polish with accent on the penultimate, the accentuated syllable represents a prominent feature of the phonological structure of lexical elements and thus a clear and distinct signal for starting the search and for the successful finding of lexical elements.
The information which is still needed at this point in order to be able to reconstruct completely the utterance containing several words will be processed and gained in the next step where verification is carried out by a component called the Master. Here, accessing the remaining information in the possible phonological structure and the top-down component, at this point especially syntax, pragmatics, and semantics, the missing parts of the phonological-syntactic structure are hypothesized and built into the total structure corresponding to (parts of) the utterance. After this verification, the process of speech recognition, hopefully, will end up with the identified meaning. As can be seen in Fig. 8, the Master has access to the linguistic constraints and the knowledge which, in turn, have access to the three lower levels. For the Master there is also a feed-back channel to the possible phonological structure which, again in turn, feeds back to the two lower levels. Thus it becomes quite clear that the top-down information is available to different and rather low levels of processing in speech recognition. It becomes also clear that, due to this fact, the speech signal need not be clear and distinct at every point in time. Of course, the more distinct the signal is, the easier and faster the lexical search can be because almost no support by the top-down component and no feeding-back is needed in this case. If the verification of some chosen lexical elements by the Master as to their linguistic and pragmatic correctness and of their semantic credability comes out negative, the feed-back channel to the possible phonological structure, the hypothesized phonological structure and, if necessary, to the acoustic-phonetic basic information will be activated. Then a change of the phonolgical structure already arrived at will be enforced by starting the searching process anew which, finally, will arrive at an acceptable result after having passed through a number of stages a second and maybe a third time.

In this interactive process of speech recognition, it is obvious that prosody, especially word accent, plays a direct and guiding part. Searching for lexical elements stored in LTM takes place not by using words with clearly defined boundaries but rather by using prosodic features where word accent and phrase accent or focus distinctly point to the most important semantic elements of an utterance. The syllables which are prominent due to word accent represent reliable islands in the stream of sounds and there they function as the anchor or fixation points of speech recognition. Therefore it is easily understood that word boundaries are not an absolute and significant support or even a precondition for speech recognition. Phrase boundaries, however, play an important part in dividing the speech chain into appropriate processing units. It is inter‿esting to notice in this respect that phrase boundaries are clearly marked, often by several prosodic means. In contrast, word boundaries, are not marked in any special way. Even where morphological word structure is concerned, unstressed syllables, especially at the end of a word, as markers of concord, normally contain linguistic information which can easily be derived. Therefore it is not astonishing to learn that speech recognition systems cannot find words in the signal of continuous speech if the word, even in longer texts, are not pronounced in a staccato way, i.e. surrounded by pauses. In the speech signal there are no word boundaries but acoustically more distinct and elaborated chunks of the size of a syllable, namely the prominent and accentuated syllables.

The model of speech perception outlined here differs from previous models in several respects, although some parts, especially at the more peripheral levels, coincide. In the present model, prosodic information in the signal and in the linguistic constraints applying to different levels and structures play a leading and guiding part in solving the task of searching for a lexical element, namely the finding and identifying of, above all, basic semantic elements, making up the skeleton of meaning.

In contrast to the cohort theory, there is no activating of groups of possible word candidates all of them beginning with the same sound and the number of which will be gradually decreased as a consequence of acoustic information arriving later and of contextual constraints until, in the end, only one candidate will hold the floor. In my model, the spectral information of phonemes does not play a predominant part. Guided by the prosodic information pointing especially to the clearly marked accentuated syllable, one or more possible phonological structures not exactly defined by word boundaries, may start for the search of lexical elements. Very often they may even act as competitors (cf. Bannert 1980).

Rather as an amendment to the Phonetic Refinement Theory, in my model the strong part of prosody in finding the most significant and central elements of meaning is duly recognized. The process of speech recognition obeys the principle of clarity. The accent pattern, prominent in the signal and easily to be discovered and processed, forms a linguistic frame or skeleton which the spectral features are subordinated to and built into. Every part of the phonological structure which is missing or indistinct, if possible, will be restored or corrected later in the interactive processes.

Another virtue of this model lies in the fact that it is applicable to the whole range of different conditions of the speech signal in verbal communication and the bottom-up component of speech perception. The top-down component is always at work. It is obvious that a distinct and good speech signal makes speech recognition easier, faster, and accurate. If the speech signal is deviant with respect to a given (band of) norm or distorted by external sources, a larger period of time will be needed in order to identify a meaning because a larger burden is put onto all kinds of memory, information paths, and feed-back channels. An increased activation of search processes and memories explains the fatigue experienced by listeners who are exposed to speech in noisy environments or to strong foreign accent for longer stretches of time.

In conclusion, then, this model also covers speech recognition under different conditions: the optimal speech signal, spoken distinctly and free from external acoustic distortions, the speaker and listener using approximately the same standard of pronunciation; the indistinct pronunciation due to lax or fast articulation; the acoustically distorted signal; the perception of the hard of hearing and the deaf; the perception under inattentiveness and non-listening of the intended listener; the geographical, dialectal, social, and individual varieties of a language; the foreign accent.

Recognizing foreign accent

There is clearly no doubt that the speech signal containing foreign accent is analysed auditorily and acoustically in the same way as the speech signal derived from standard language. First differences are to be found at the point of the acoustic-phonetic analysis. Searching for lexical elements cannot be done in real time, because the incomplete and fragmentary basic acoustic-phonetic information does not permit generating a hypothesized phonological structure leading to a possible phonological structure. As a consequence of this failure, information has to be kept in the short-term memory which puts an extra load on it, while the searching for a word is expanded by waiting for more phonetic bottom-up information and by switching on the top-down restoration and corrections components. This, in turn, will put even more strain on the recognition processes.

Another problem for lexical search arises when the possible linguistic structure points to the wrong lexical element. This is the case when a word pronounced deviatingly coincides with a different, existing word; for instance when the phoneme /y/ is rendered as the phoneme /i/ (Swedish byta - bita 'change - bite'). In this case, the lexical search seemingly will succeed in identifying a word and finding a meaning. However, this mistake will be discovered when the word is put into the phrase or sentence where the context discloses that the wrong word was picked. The interpretation of the whole phrase or sentence has to be rejected at this stage and a new recognition process has to be started, now also by activating the restoration component. Again a greater strain is put on the processing of the speech signal. Furthermore, it has to be pointed out that the speech signal, while repetitions and retentions in the short-term memory are in full progress, continues to enter the ear, and the peripheral automatic acoustic analysis must continue its work without interruption.

The decoding processes for foreign accent should show the heaviest strain with listeners who are not accustomed to this phonological variation and who are not motivated to do such extra labour. The decoding processes for foreign accent should show the lightest strain with listeners who have developed in their long-term memory a rich component of correction rules for foreign accent - which is closely related to the typical features of foreign accent of a given L1 - and who really want to understand foreigners by activating both the feed-back path (the correction component) and the access path of top-down information (the restoration component).

FOOTNOTES

<*> I would like to thank Klaus-Jürgen Engelberg, David House, Bernhard Keck, Gerhard Rigoll, and Herbert Tropf for helpful and valuable comments and contributions to this paper.
This research was supported by the Bank of Sweden Tercentenary Foundation.

<1> The manipulations of the speech signal were made at the Department of Linguistics, Uppsala University. I am very grateful to Sven Öhman for his kind support and Lennart Nordstrand for his expert assistance.

<2> It became evident from preliminary tests using filtering as a means of distorting the acoustic speech signal that intelligibility would not be decreased to a sufficient and desired degree. Therefore the planned experiment with filtered speech was excluded (cf. Bannert 1984).

<3> The listening tests were disguised as reaction tests attempting to measure the listeners' ability, as quickly as possible, to prompt the utterances spoken by foreigners and presented under hard listening conditions.

<4> One utterance (L1 = Greek) that was part of the test was excluded from the presentation of the results as the manipulation of voicing and voicelessness of the obstruent cluster by LPC-synthesis is not reliable in this respect.

<5> Quantity in Standard Swedish is manifested as the complementary length pattern /V:C/ vs /VC:/ in the stressed syllable. Phrase rhythm means the temporal relationships between successive syllables. Accent is a tonal feature of syllable prominence and is manifested as a change of Fo in or in connection with the accentuated syllable.

<6> The parameter of volume (intensity) of syllables was not included in the manipulations of accent. This does not mean, of course, that intensity might not be a contributing factor in the complex feature of accent. It is believed, however, that intensity is not an essential feature of normal word accent (as opposed to contrast or emphasis).

<7> These prosodic deteriorations corresponding to the

features of phrase rhythm, quantity, and pitch accent, had a detrimental effect on the identity of the utterances. In several instances of demonstrations where the deteriorated Swedish utterances were played to linguists, phoneticians, and experienced teachers of Swedish as a second language, these disguised utterances were accepted as foreign accent and associated with certain first languages L1.

<8> As individuals, listeners react differently to the presentation of the stimuli. While some of them always try to respond even guessing to some degree, others hesitate to respond at all if they are not quite sure about the intended structure.

<9> For these larger units no definition is provided here. Yet it is assumed that the notion is well-established.

<10> Some possible elaborations in certain respects may look like parts of the model in Lea et al. (1975).

<11> The dimensions of voice quality and volume will also be analysed on this level. This processes are only mentioned to complete the picture and will not be dealt with here.

REFERENCES

Bannert R. 1975. The significance of vowel features in the perception of complementary length in Central Bavarian. Lund University, Department of General Linguistics and Phonetics, Working Papers 12, 1-45.

Bannert R. 1978. Prosodiska egenskapers effekt pa förståelig-heten. In: Kommunikativ kompetens och fackspråk. Linnarud M. and Svartvik J. (eds.). Lund

Bannert R. 1980. Phonological Strategies in the Second Language Learning of Swedish Prosody. PHONOLOGICA 1980, 29-33. Innsbruck

Bannert R. 1984. Prosody and intelligibility of Swedish spoken with a foreign accent. Nordic Prosody III, C-C.Elert,.Johansson, and E.Strangert (eds.). Acta Universitatis Umensis, Umeå Studies in the Humanities

59, 7-18

Bannert R., Engstrand O., Eriksson H., and Nordstrand L. 1982. Phonetics and Identity. Uppsala University, Department of Linguistics, Report 8, 22-58

Bruce G. 1977. Swedish Word Accents in Sentence Perspective. Travaux de l'Institut de Linguistique de Lund XII. Lund

Cole R.A. and Jakimik J. 1980. A Model of Speech Perception. In: Perception and Production of Fluent Speech, 133-163. Cole R.A. (ed.). Hillsdale

Elman J.L. and McClelland J.L. 1984. Speech Perception as a Cognitive Process: The Interactive Activation Model. In: Speech and Language: Advances in Basic Research and Practice, Vol. 10, 337-374. Lass N.J. (ed.). New York.

Elman J.L. and McClelland J.L. 1985. Exploiting Lawful Variability in the Speech Wave. In: Variability and Invariance of Speech Processes, Perkell J.S. and Klatt D.H. (eds.). Hillsdale, N.J.

Forster K.I. 1979. Levels of Processing and the Structure of the Language Processor. In: Sentence Processing: Psycholingustic Studies Presented to Merrill Garrett, Cooper W.E. and Walker E.T. (eds.). Hillsdale, N.J.

Gårding E. and Bruce G. 1981. A Presentation of the Lund Model for Swedish Intonation. Lund University, Department of Linguistics and Phonetics, Working Papers 21, 69-75

House D. 1985. Sentence prosody and syntax in speech perception. Lund University, Department of Linguistics and Phonetics. Working Papers 28, 91-107

Klatt D.H. 1979. Speech Perception: A Model of Acoustic-Phonetic Analysis and Lexical Access. J. of Phonetics 7, 279-312.

Klatt D.H. 1980. Speech Perception: A Model of Acoustic-Phonetic Analysis and Lexical Access. In: Perception and Production of Fluent Speech, 242-288. Cole R.A. (ed.). Hillsdale, N.J.

Lea W.A. 1980. Prosodic Aids to Speech Recognition. In: Trends in Speech Recognition , 166-205. W.A.Lea (ed).

Englewood Cliffs, N.J.

Lea W.A., Medress m.F., and Skinner T.E. 1975. A Prosodically Guided Speech Understanding Strategy. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-23, 30-38

Maassen B. and Povel D-J. 1985. The effect of segmental and suprasegmental corrections on the intelligibility of deaf speech. JASA 78, 877-886

Marslen-Wilson W.D. and Welsh A. 1978. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology 10, 29-63

Morton J. 1979. Word Recognition. In: Psycholinguistics 2: Structures and Processes, 107-156. Morton J. and Marshall J.C. (eds.). Cambridge, MIT Press.

Nespor M. and Vogel I. 1983. Prosodic Structure Above the Word. In: Cutler A. and Ladd D.R. (eds), Prosody: Models and Measurements, 123-140 (=Language and Communication 14). Berlin

Nooteboom S.G. 1981. Lexical retrieval from fragments of spoken Words: Beginnings vs. endings. Journal of Phonetics 9,407-424

Osberger M.J. and Levitt H. 1978. The effect of time errors on the intelligibility of deaf children's speech. JASA 66, 1316-1324

Pisoni D.B. 1984. Acoustic-Phonetic Representations in Word Recognition. Indiana University. Research on Speech Perception, Progress Report No.10, 129-152

Pisoni D.B., Nusbaum H.C., Luce P.A., and Slowiaczek L.M. 1985. Speech Perception, Word Recognition and the Structure of the Lexicon. Speech Communication 4, 75-95

Sendlmeier W.F. 1985. Psychophonetische Aspekte der Wortwahrnehmung (= Forum Phoneticum 35). Hamburg

Shipman D.W. and Zue V.W. 1982. Properties of Large Lexicons: Implications for Advanced Isolated Word Recognition Systems. ICASSP 82, 546-549.

Svensson S-G. 1974. Prosody and Grammar in Speech Perception. University of Stockholm, Monographs from the Institute of Linguistics 2

van Wijk C. and Kempen G. 1984. From sentence structure to intonation contour. Germanistische Linguistik 79-80, 157-182. Müller B.S. (ed.). Hildesheim

Independence and interdependence of prosodic features

Robert Bannert

ABSTRACT

Considering existing prosody models, two fundamentally different approaches can be discerned. In one case, the tonal structure is treated as independent, whereas the temporal structure is seen as dependent and derivable from the tonal structure. This approach assumes the primacy of intonation. In the other case, the two dimensions of time (duration) and frequency (Fo) are treated as independent of each other. Therefore the one cannot be derived from the other. Instead the basic temporal and tonal structures are generated separately. This approach ascribes time and intonation an autonomous status.

Starting from this dichotomy, this paper will promote the discussion about the principles of building prosody models. It seems essential to abandon the categorical question about the either-or status of time and intonation and to recognize the complex interrelationships between these two dimensions. Therefore time and intonation should be considered equal in principle, although it is quite obvious that there exist certain relationships between them.

An attempt is made to illustrate the approach of equality between duration and Fo using Swedish test material. Aspects of word and sentence level prosody are investigated. The independence and interdependence of duration and Fo will be displayed. The question which is put forward and which seems more fruitful is not whether there are any dependencies but rather what the interrelationships look like. It will also be demonstrated how tonal features behave in a case of extreme time shortage. When several tonal features of an utterance are forced into one single syllable, a total reorganization of the tonal contour is to be observed exhibiting a clear tonal hierarchy on the word and sentence level.

The observations on the Swedish material are supported by references to equivalent phenomena in some other languages thus lending a more general character to them.

The results of this investigation are the starting point for the outline of a new prosody model. The tonal and temporal structures of utterances will now be generated in parallel with interactive processes. Linguistic rules and information of different kinds are applied. Therefore the adjustment component in an earlier version of the model is disposed of.

Last not least, shedding light on the relationships between time and intonation is important also for the development of high-quality speech synthesis in text-to-speech systems.

INTRODUCTION

For over ten years now, a discussion has been going on concerning the relationship between the two prosodic features of segment duration, i.e. the temporal structure of utterances, and the tonal movements in utterances, i.e. their tonal structure. The question has been whether these two features are independent of each other or if one can be derived from the other. Adherents of the latter view assume that the tonal gestures (movements) constitute the primary, basic feature out of which the segment durations follow as an automatic consequence of the tonal demands and requirements. This stand, which may be termed the primacy of Fo in a prosody and speech model, is taken, for instance, by Öhman et al. (1979) and Lyberg (1981).

Opposing this view, the time and tone dimensions of speech are considered to be separate entities, each of which exists on its own grounds. However, time and frequency do not exist independently of each other. Nevertheless, in a generative prosody model, the basic temporal and tonal structures of an utterance are indeed generated separately of each other. The temporal structure is processed first, because it serves the tonal structure, defined by its tonal anchor points <1>, as a reference for projection. Then the basic temporal and tonal structures are added where different kinds of adjustments become necessary. This is the case when a tonal gesture or successive tonal gestures only have a limited time to be executed. The resulting tonal conflicts are of two kinds: time-dependent and position-dependent (Bruce 1977, 74). The approach which considers time and frequency as separate dimensions, although time is seen as primary delimitating frequency in cases of conflict between them, is represented by Thorsen (1980), Bruce (1977, 1981), Gårding et al. (1982), and Bannert (1982a,b) and may be termed the autonomous model of prosody.

A discussion of the relationships between tonal and temporal features in a prosody model and a first examination of Lyberg's model of Fo-dependent segment duration is to be found in Bannert (1982a).

Taking these opposing approaches as the starting point, it is the aim of this paper <*> to continue the discussion and to arrive at a clearer picture of the principles of a prosody model <2>. It will be asked if it is justified at all to formulate categorical questions about the dependence or independence of time and intonation since data suggest that there is a complex acting together of segment durations and

Fo. Therefore the dimensions of time and frequency should be treated as equal partners and processed separately, although they share independencies and interdependencies. Using Swedish material which contains temporal and tonal variations, these interrelationships will be demonstrated.

Compared to previous studies, the present investigation also widens the number of variables by including the following three variables: (1) the opposite tonal manifestation of identical tonal features (word accent II and sentence accent) in two Swedish dialects (Standard and Southern Swedish), (2) the quantity (complementary length of the stressed vowel and the following consonant in Standard Swedish and long/short vowel contrast in Southern Swedish), and (3) three, different, non-final sentence positions of the test word.

THE INVESTIGATION

The variables are presented that are used for the intended variation of time and frequency. Then the design of the test is shown and information about the recordings and the analysis is given.

Variables

The following variables were changed in a statement spoken as the answer to an appropriate question:

sentence accent
quantity
sentence position of test word (sentence medial)
dialect (Standard Swedish, Southern Swedish)
speakers

For the manifestation of the prosodic features the following differences can be observed:

Besides the tonal differences in the accentuated vowel of word accent II in both dialects (a fall in Standard Swedish, a rise in Southern Swedish), sentence accent is manifested strikingly differently.

There are also dialectal differences as to the manifestation of quantity. Whereas the stressed VC-sequences show the pattern of complementary length (/V:C/ vs /VC:/) in Standard Swedish, Southern Swedish displays quantity in the stressed vowel only (/V:C/ vs /VC /).

Material

As the starting material, the following sentence was chosen which, with respect to its phonetic and syntactic structure, corresponds to a well-established standard in intonation studies of Swedish:

Man	kan	l\`ämna	l\`ånga	n\`unnor	efter	\`åtta.
		1	2	3		
		verb	adject-ive	noun		

(You can leave long nuns after eight o'clock)

\` = word accent II (grave accent)
1, 2, 3 = position for test words

Test words were stöka with a long vowel and stöcka with a short vowel which were also used in Bannert (1979). The two test words were inserted in turn into the three sentence positions. Sentence accent was placed on the three positions using questions as appropriate contexts. Otherwise, when the test words should not be in focus, sentence accent was placed on the time adverbial (åtta) at the end of the sentence. Thus it was ensured that the word accents were not influenced by the sentence accent because the word accent in position 3 was followed by three unstressed syllables preceding the final word carrying sentence accent. In all, the whole material consisted of twelve sentences: six sentences where the test words did not carry sentence accent, the time adverbial being focussed, and six sentences with sentence accent on each of the three positions and the two test words. Sentence accent was shifted by asking questions about the test words in the different positions (cf. the method used in Bruce 1977, 21 ff.).

Recordings and analysis

The test material was read in a kind of one-person dialogue of question and answer (= test sentence) by four speakers seven times each. Informants were TB (male) and EH (female) from Stockholm (identical with the informants in Bannert

1979) representing Standard Swedish and EK and AO (both female) from Malmö and Lund, respectively, representing Southern Swedish. The sentences were read fluently as one single prosodic phrase, i.e. they were produced in one breath without pausing before the time adverbial. The material was recorded in the acoustic studio of the Department of Linguistics and Phonetics, Lund University, using a STUDER-tape recorder A 62 at the speed of 7.5 ips. The recordings were analysed acoustically using a Frøkjaer-Jensen Pitch Meter yielding a duplex oscillogramme and an Fo curve and, at the same time, a FONEMA Intensity Meter yielding an intensity curve; recording speed was 100 mm/s. The registrations were segmented by hand; segment durations (a[n], s, t, ö, k, a) were measured with an accuracy of 5 ms; Fo was measured at four points A, B, C, D defined below (cf. Fig. 1) with an accuracy of 5 Hz. The individual means were calculated and rounded off to the nearest 1 ms and 1 Hz respectively. Standard deviations were also calculated.

RESULTS

Superimposed tonal contours of a typical utterance containing the test word <u>stöka</u> as an adjective in position 2 with and without sentence accent for both dialects are shown in Fig. 1. The four points of tonal measurement A, B, C, and D are indicated and defined as follows:

- Point A: End of the preaccentuated, unstressed vowel.
- Point B: Beginning of the accentuated vowel. The Fo-minimum in the Southern Swedish curves are most often preceded by a short, small fall.
- Point C: End of the accentuated vowel. The Fo-maximum in the Southern Swedish curves are most often followed by a short, small fall.
- Point D: Fo-value at the VC-boundary, i.e. at the end of the unstressed vowel [a] in the second syllable of the test word.

The results are presented as follows:
First, based on the means of all the measured values of segment durations and Fo in the four tonal points, some general observations are made. An overview of the tonal aspects of the material is given in Fig. 2 where the Fo-points are plotted time normalized including all variables for each speaker. Then, based on the mean values, the influence and effect of quantity, sentence accent, and

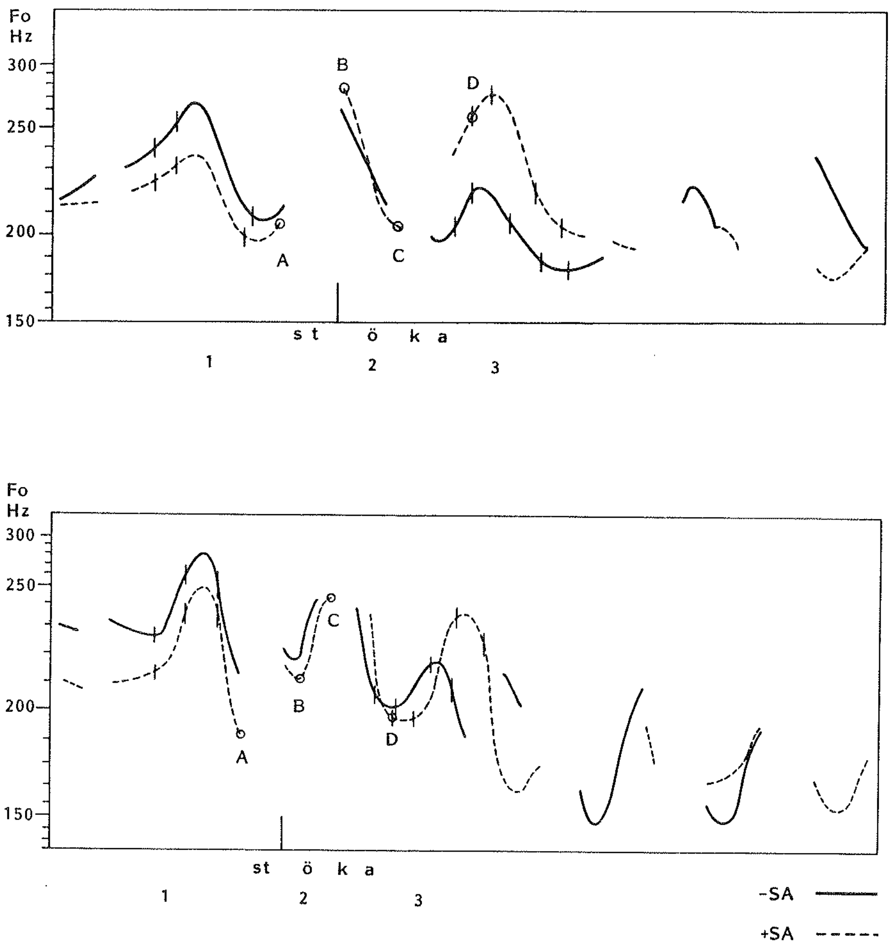


Fig. 1 Superimposed, typical Fo-contours of the test word in medial sentence position (position 2) with and without sentence accent. Four points of tonal reference (A, B, C, D) are shown. Standard Swedish, speaker EH, above; Southern Swedish, speaker EK, below.

sentence position on the temporal and tonal structure of the test words respectively are reported. In each case, the differences calculated from the means are given in tables. Finally a conflict situation between time and frequency is shown where time dominates over frequency, and a tonal hierarchy at work is illustrated.

Segment durations and Fo-values

On the basis of mean values, the following general observations can be made:

Duration

1. Sentence accent increases the duration of all segments with all four speakers, although in some cases only to a small extent. There are also instances, however, where the increase is considerable. There is only one exception: Speaker AO, final [a], all positions, where a systematic decrease of segment duration is to be found.

2. Many segments show smaller durations in sentence position 2 compared to the other positions. This seems to be a positional effect, i.e. an expression of the rhythmical organization. According to this principle, a succession of two or more equal accents is avoided by weakening the accent in the middle temporally as well as tonally (cf. Bruce 1983 for Swedish and Bannert 1983 for German).

3. The VC-sequences of the Stockholm speakers show the typical pattern of complementary length (cf. for instance Elert 1964) which the Southern Swedish speakers do not have. Their consonants following short vowels are only slightly longer (cf. Gårding et al. 1974).

Fo

Fig. 2 clearly shows the different tonal behaviour of the Fo-points and the Fo-movements with and without sentence accent, with long and short vowels, in the three sentence positions, between the two dialects, and between the two speakers of each dialect. It should be noted, however, when inspecting the curves that only the movement between point B and C (beginning and end of the accentuated vowel) is to be seen completely in the registrations (cf. Fig. 1). The other two points are simply connected by straight lines. For the

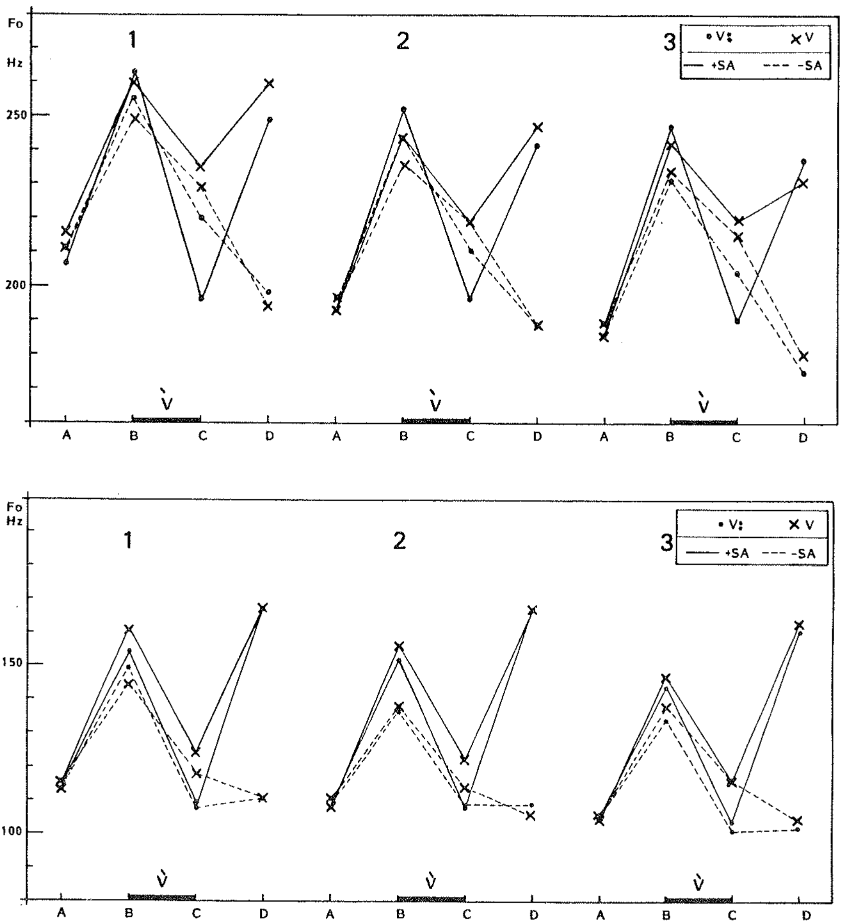


Fig. 2a. Superimposed Fo-contours (time normalized) of the test word in the three positions 1, 2, 3 for the Standard Swedish speakers (EH above, TB below). Four conditions: Long/short vowel, with and without sentence accent.

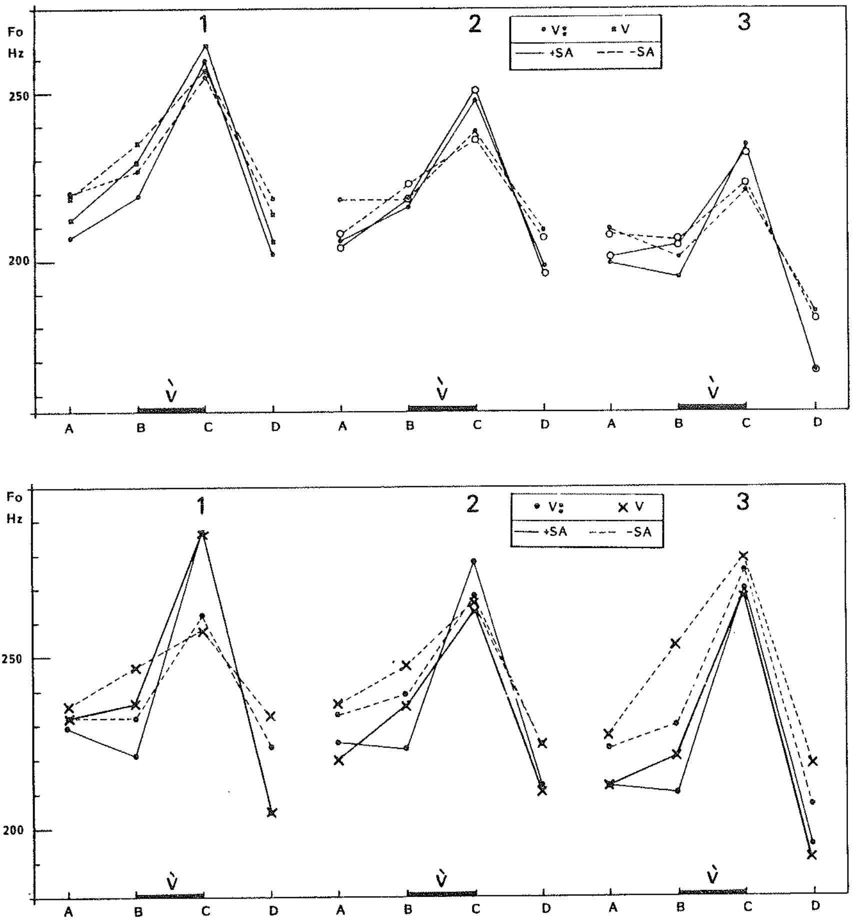


Fig. 2b. Superimposed Fo-contours (time normalized) of the test word in the three positions 1, 2, 3 for the Southern Swedish speakers (EK above, AO below). Four conditions: Long/short vowel, with and without sentence accent.

aim of this investigation, the variation of the Fo-points is of interest and not the complete Fo-movement.

All variables show an effect on the Fo-values, although the shape and movement of the tonal contours, by and large, are preserved in each case. The tonal gesture of the word accent is treated differently in the two dialects. Whereas the tonal fall of accent II in Standard Swedish is truncated, the tonal rise in Southern Swedish is reorganized (cf. Bannert and Bredvad-Jensen 1975). Sentence accent shows up tonally not only in the post-accentuated syllable - as a high point in Standard Swedish and as a low point in Southern Swedish - but also in the accentuated syllable itself (this is clearly to be seen in the curves of the Southern Swedish speakers). Thus sentence accent exerts an influence tonally and temporally on the whole test word. This influence, though, is still greater where the pre-focal accent and the overall shape of the sentence contour is concerned (this effect can be clearly seen in Fig. 1).

Quantity and sentence position also affect the tonal structure. In Fig. 2, the Fo-declination throughout the utterance is to be seen exhibiting differnces between the speakers.

What, then, are the effects in this particular case that quantity, sentence accent and sentence position have on the tonal and temporal structure of the test word in both dialects? We will look for patterns of variation or consistency that can be found either in the whole material or in one dialect, respectively.

Quantity

Durations

The durational changes of all test segments as a consequence of quantity are calculated and given in Table 1. A minus sign indicates that the duration of a given segment is larger in the word with the short vowel. As for the following tables, the values in Table 1 are derived from the means of the basic data. For the sake of simplicity, the standard deviations are omitted <3>.

Table 1 shows that, as a rule, the duration of all segments varies in all conditions. The largest durational differences

Table 1. Differences of segment durations (ms) due to quantity.

POSITION						1								2								3				
SEGMENTS			an	s	t	ö	k	a	VC	WORD	a	s	t	ö	k	a	VC	WORD	a	s	t	ö	k	a	VC	WORD
STOCKHOLM																										
	TB	+SA	2	5	2	79	-41	0	38	43	1	1	0	79	-54	15	25	26	0	1	-3	75	-55	-7	20	10
		-SA	-4	-3	3	33	-21	-5	12	4	3	10	0	30	-18	-7	12	10	5	2	-5	43	-18	-2	25	13
	EH	+SA	-5	6	-5	50	-46	-14	4	-7	1	0	8	50	-34	3	16	27	3	9	5	63	-39	-9	24	29
		-SA	1	8	-4	29	-20	-11	9	0	-1	1	4	25	-24	-3	1	16	2	1	2	36	-21	-5	15	8
SKÅNE																										
	EK	+SA	-1	-1	-10	34	-15	-3	19	5	-3	-8	-8	32	-14	-4	18	-2	1	-7	-5	42	-15	3	27	15
		-SA	2	-1	-7	35	-11	-7	24	11	-4	-8	-2	28	-9	-6	19	3	-2	-2	-10	35	-13	-10	22	1
	AO	+SA	9	0	-13	57	-38	-2	19	2	4	-2	-10	61	-12	-6	49	33	3	1	-10	65	-21	3	44	37
		-SA	6	15	-5	44	-5	-3	39	42	-1	8	-11	47	-9	-5	38	30	2	-4	-2	45	-12	-9	33	13

Negative values indicate that the segment durations with the short vowel are larger than with the long vowel.

are to be found in the accentuated vowel and the following consonant, especially in the VC-sequence with sentence accent. Most clearly, this difference is to be seen with the Stockholm speakers. Thus the VC-sequences display a consistent pattern of variation. The other segment durations vary, by and large, only to a small extent and without any discernible pattern.

<u>Fo-points</u>

Table 2 shows the differences of the Fo-means in the four tonal points A, B, C, and D as a consequence of quantity (cf. Fig. 2). The minus sign indicates that the Fo-value in the test word containing the short vowel is greater (the point is higher) than in the test word with the long vowel (cf. Fig. 1). In this case, except for a few instances especially in sentence position 2, a consistent pattern of variation for the whole material is to be found. The Stockholm speakers show a tonal difference at the end of the vowel (Fo-point C) which is considerably larger than at the beginning. The end point of the short vowel contour with and without sentence accent in each position is clearly higher than that of the long vowel contour. The Southern Swedish speakers show the largest tonal difference at the beginning of the vowel (Fo-point B), the short vowel causing the highest Fo-values. The Fo-values in the other points A and D, in general, vary only slightly and inconsistently.

Sentence accent

<u>Durations</u>

With a few exceptions, sentence accent causes an increase in the segment durations. Table 3 gives the differences of segment durations as a consequence of sentence accent. The minus sign indicates that the segment duration in the test word without sentence accent is larger than in the test word with sentence accent. Large and systematic variations are to be found in the segments [s], the accentuated vowel [o], the following consonant [k], and the unstressed vowel [a]. The smallest and most inconsistent durational changes are to be observed in the segments [a(n)] and [t]. In Standard Swedish,

Table 2. Differences of Fo (Hz) at the four Fo-points due to quantity.

POSITION			1				2				3			
FO-POINTS			A	B	C	D	A	B	C	D	A	B	C	D
STOCKHOLM		+SA	-1	-7	-15	0	1	-4	-14	1	1	-3	-12	-3
	TB	-SA	-3	5	-10	0	-1	-1	-5	3	0	-4	-15	-2
		+SA	-8	3	-37	-10	1	8	-23	-5	1	5	-29	6
	EH	-SA	0	6	-9	4	1	8	-8	0	-1	-3	-11	-4
SKANE		+SA	-5	-11	-5	-4	2	-2	-3	2	-2	-10	1	0
	EK	-SA	1	-8	-2	5	10	-5	2	2	1	-5	-2	2
		+SA	-3	-15	1	0	5	-13	14	3	0	-11	2	4
	AO	-SA	-3	-15	4	-8	-3	-8	2	0	-4	-23	-4	-10

Negative values indicate that the Fo-points with short vowels are higher than with long vowels.

Table 4. Differences of Fo (Hz) at the four Fo-points due to sentence accent.

POSITION			1				2				3			
FO-POINTS			A	B	C	D	A	B	C	D	A	B	C	D
STOCKHOLM		V:	1	5	1	57	0	15	-1	58	1	10	7	59
	TB	V	0	17	6	57	-2	18	8	61	0	9	0	60
		V:	-4	8	-24	51	-2	12	-15	53	3	16	-14	62
	EH	V	4	11	6	65	-2	8	0	58	-3	8	4	52
SKANE		V:	-13	-8	5	-17	-12	-2	10	-11	-10	-6	13	-18
	EK	V	-7	-5	12	-8	-4	-5	15	-11	-7	-1	10	-16
		V:	-3	-11	26	-19	-8	-16	10	-12	-11	-20	-5	-12
	AO	V	-3	-11	29	-27	-15	-11	-2	-15	-15	-32	-11	-26

Negative values indicate that the Fo-points with sentence accent is lower than that without sentence accent.

Table 3. Differences of segment durations (ms) due to sentence accent.

POSITION						1								2								3				
SEGMENTS			an	s	t	ö	k	a	VC	WORD	a	s	t	ö	k	a	VC	WORD	a	s	t	ö	k	a	VC	WORD
STOCKHOLM																										
	TB	V:	5	22	15	56	27	23	83	143	1	16	12	61	26	19	87	130	1	16	9	43	31	26	74	135
		V	-1	14	16	10	47	16	57	104	3	25	12	12	62	7	74	114	6	17	7	11	68	31	79	138
	EH	V:	8	18	-1	35	16	8	51	76	2	19	4	43	20	11	63	96	8	19	3	44	17	13	61	96
		V	12	20	0	14	42	11	56	83	0	20	0	17	30	5	47	73	5	11	0	17	35	19	52	77
SKÅNE																										
	EK	V:	-1	11	2	13	9	5	22	36	-7	9	-3	18	7	2	25	32	2	16	3	23	18	42	41	101
		V	2	11	5	14	13	1	27	43	-8	9	3	14	12	0	26	37	-1	21	-2	16	20	29	36	87
	AO	V:	2	11	1	27	31	-1	58	73	-3	-3	5	17	15	-3	32	30	1	19	1	27	25	-7	52	69
		V	-1	26	9	14	64	-2	78	113	-8	7	4	3	18	-2	21	26	0	14	9	7	34	-19	41	45

Negative values indicate that the segment durations without sentence accent are larger than those with sentence accent.

the durational differences are largest in the long segment (V: and C:), respectively.

Fo-points

Sentence accent also affects the Fo-values. Table 4 shows the Fo-differences in the four tonal points A, B, C, and D as a consequence of sentence accent. The minus sign indicates that the value of the Fo-point is larger in the test word without sentence accent, i.e. it is higher. No consistent pattern for the whole material can be found. Within the two dialects, however, there is a similar tonal behaviour.

With the Stockholm speakers, the large tonal difference appears in point D. At this point (VC-boundary), the Fo-maximum of the sentence accent is almost reached. The high tonal point at the beginning of the word accent fall (point B) is also higher with the sentence accent, both the long and short vowel and in each sentence position. Point A, the Fo-minimum in the pre-accentuated syllable remains nearly unchanged. At point C, the Fo-minimum of the word accent fall, the speakers behave differently. Whereas speaker TB hardly varies in this point, speaker EH makes the word accent fall with sentence accent end considerably lower only in the long vowel.

The picture is more uniform with the Southern Swedish speakers. The Fo-points A, B, and D are lower with sentence accent, point C is higher (one exception: speaker AO, position 2, short vowel and position 3). This means that the tonal movement before and in the test word is larger with sentence accent, i.e. the Fo-curve makes a larger excursion up and down, the tonal movement shows a larger range (cf. Fig. 2).

Positions

Durations

Segment durations vary also as a consequence of sentence position. Table 5 gives the durational differences between the positions where the value of position 2 serves as a reference. The first value in Table 5 corresponds to the

Table 5. Differences of segment durations (ms) between positions.

SEGMENTS				a(n)	s	t	ö	k	a	VC	WORD
STOCKHOLM											
	TB	V:	+SA	18	1	6	0	3	-9	3	5
				5	1	-6	0	9	8	9	16
			-SA	14	-5	3	5	2	-5	7	-8
				5	1	-3	18	4	1	22	11
		V	+SA	17	-3	4	0	-10	-4	-10	-11
				6	1	-3	4	10	20	14	32
			-SA	21	8	0	2	5	-7	7	-2
				3	9	2	5	4	-4	9	8
	EH	V:	+SA	12	18	-11	8	1	-13	9	0
				6	11	-1	16	4	-3	20	28
			-SA	6	19	-6	15	5	-10	20	20
				0	11	0	14	7	-5	21	28
		V	+SA	18	12	2	8	13	4	21	34
				4	2	2	3	9	11	12	26
			-SA	4	12	2	11	1	-2	12	24
				-3	11	2	3	4	-3	7	22
SKÅNE	EK	V:	+SA	-2	6	-1	9	0	3	9	16
				12	15	4	20	24	49	44	111
			-SA	-8	4	-6	14	-2	0	12	11
				3	8	-2	15	13	9	28	44
		V	+SA	-4	-1	1	7	1	2	8	9
				8	14	1	10	25	44	35	94
			-SA	-14	-3	-1	7	0	1	7	3
				1	2	6	8	17	13	25	44
	AO	V:	+SA	14	4	-10	6	8	2	14	22
				14	3	-9	12	20	13	32	54
			-SA	9	2	-6	-4	-8	0	-12	-20
				10	-5	-5	2	10	17	12	16
		V	+SA	9	14	-7	10	34	-2	44	53
				15	14	-9	8	29	4	37	50
			-SA	2	-5	-12	-1	-12	-2	-13	-33
				7	7	-14	4	13	21	17	33

Line above: difference between 1st and 2nd positions, negative value indicates that the segment duration in the 1st position is smaller than that in the 2nd position

Line below: difference between 2nd and 3rd position, negative value indicates that the segment duration in the 2nd position is smaller than that in the 3rd position durational difference of a given segment between position 1 and 2, the second value to that between position 2 and 3. A negative value indicates that the segment duration in the first position is smaller than that in the second position and that the segment duration in the second position is smaller than that in the third position <4>. As Table 5 shows, no pattern of variation of segment duration between positions can be found, neither for the whole material, nor for each dialect, nor for each speaker. This is also true of the duration of the VC-sequence and the word.

Fo-points

With only a few exceptions, above all for speaker AO, the position of the test word in the sentence, i.e. its placement in the tonal contour of the sentence with reference to the time axis, affects the Fo-points in a systematic way. Table 6 gives the differences in Hz of the four tonal points A, B, C, and D between the sentence positions (cf. Fig. 2). The first value in Table 6 is the Fo-difference between positions 1 and 2, the second value is the Fo-difference between positions 2 and 3. A minus sign preceding the first value indicates that the Fo-value in position 2 is larger than that in position 1. A minus sign preceding the second value means that the Fo-value in position 3 is larger than that in position 2. Table 6 shows that each Fo-point decreases the further it is located to the right in an utterance. Thus all the four points obey the following rank order: 1 < 2 < 3, i.e. the Fo-values are largest in position 1 and smallest in position 3. In other respects, however, no systematic variation is to be observed. Nevertheless, there are individual differences as to the size of the tonal differences according to position. Whereas the Fo-values of speaker TB only show small differences, they drop, sometimes considerably, with the other speakers. This means that the Fo-declination of speaker TB is rather small, the other speakers showing a clear declination (cf. also Fig. 2). Speaker AO manifests the largest irregularities in her tonal variation.

The two-dimensional analysis of this investigation has shown that the tonal and temporal features affect each other mutually to some degree. However, the prosodic features themselves are preserved as characteristic temporal and tonal patterns. Sentence accent turns out to be a prosodic feature with a Janus face; it is clearly signalled tonally as well as temporally over the whole word which it makes prominent on

Table 6. Differences of Fo-points (Hz) between positions.

FO-POINTS			A		B		C		D	
DIFFERENCE BETWEEN POSITIONS			1-2	2-3	1-2	2-3	1-2	2-3	1-2	2-3
STOCKHOLM										
TB	V:	+SA	5	3	2	8	1	4	0	7
		-SA	4	4	12	3	-1	8	2	7
	V	+SA	7	3	5	9	2	6	1	3
		-SA	5	5	6	0	4	-2	5	2
EH	V:	+SA	14	4	11	5	0	6	7	5
		-SA	16	9	11	13	9	7	9	14
	V	+SA	23	6	16	2	16	0	12	16
		-SA	17	5	13	2	10	4	5	10
SKANE										
EK	V:	+SA	1	7	3	21	12	14	4	31
		-SA	2	9	9	17	17	17	10	24
	V	+SA	8	3	12	13	14	18	10	29
		-SA	11	0	12	17	19	13	7	24
AO	V:	+SA	4	13	-2	13	10	8	-7	19
		-SA	-1	10	-7	9	-6	-7	0	19
	V	+SA	12	8	0	15	23	-4	-4	20
		-SA	0	8	0	-6	-8	-13	8	9

the sentence level.

In the present investigation, the tonal features in each position had enough time (syllables) at their disposal in order to be manifested without difficulty. However, a positive statement about the independence or dependence of variables can best be made when a conflict arises between the dimensions. Such a case of conflict between tonal features and the temporal structure of the utterance, i.e. between frequency and time, will be demonstrated in the following section.

A tonal hierarchy

Starting from a phrase with four syllables and the three tonal features of word accent I, phrase (sentence) accent and terminal juncture (statement), a stepwise reduction of the number of syllables will show the dependence of tonal features on time. The tonal feature, with the largest domain will dominate over the other features which rank lower in the tonal hierarchy.

Consider the four Swedish phrases:

å 'länderna	"and the countries"
å 'länder	"and countries"
å 'land	"and country"
'land	"country"

As the number of syllables is decreased, the duration of the phrase is also decreased, namely from about 750 ms to about 400 ms. The number of the tonal features, however, is kept constant. Thus the one-syllable utterance is produced with the same tonal features as the four-syllable utterance. Fig. 3 shows the changes of time and Fo-contour in each utterance spoken by a Stockholm speaker. In this time-dependent case of conflict, it becomes obvious that the complex overall Fo-contour shrinks considerably. The tonal contour is reorganized displaying a clear hierarchy between the three tonal gestures each of which can be seen in the complete tonal contour (cf. also Bruce 1977, 92 ff.). The final fall associated with the terminal juncture is preserved in the one syllable phrase but, at the same time, it becomes steeper and "moves" to the left on the time axis into the final (and only) accentuated vowel. The tonal movements associated with the two other gestures are skipped. The subordination of

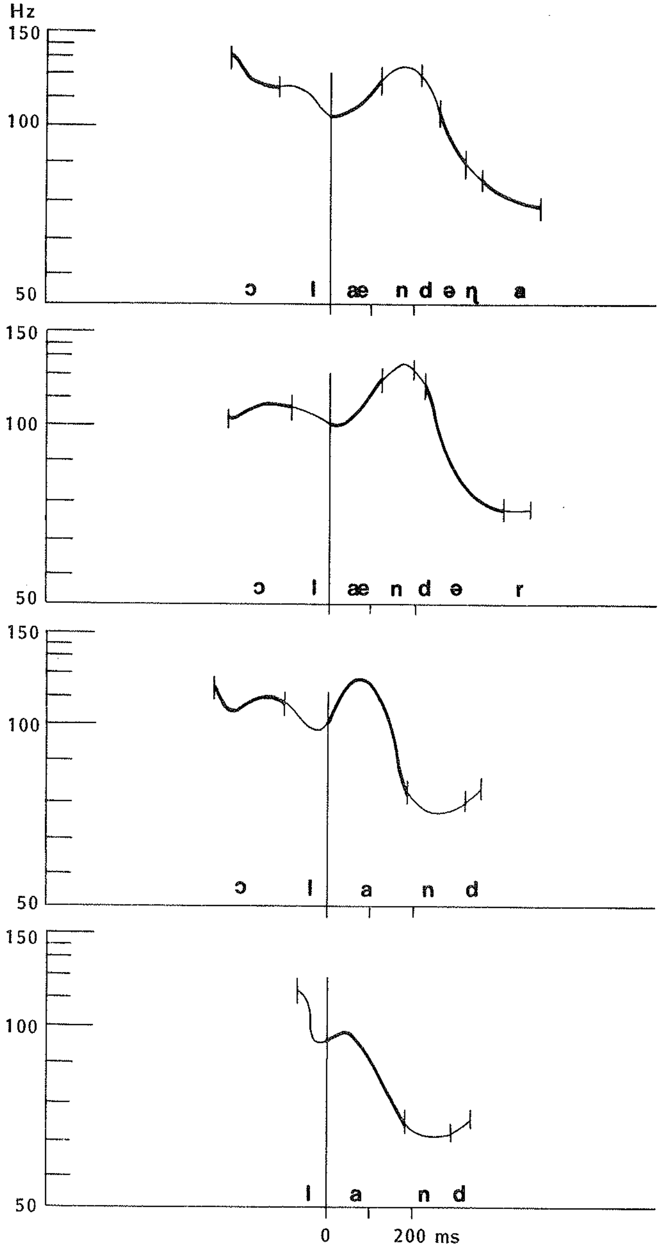


Fig. 3 Hierarchical order of tonal features illustrated by a step-by-step shortening of the duration available for three tonal features. The final fall of the terminal juncture dominates over the features of phrase accent and word accent which have smaller domains.

tonal features of a temporally smaller domain under the feature of terminal juncture with the largest domain is to be observed in other languages too. Even if the hierarchically lower gestures are extinguished from the Fo-curve, they, however, do not disappear for the listener. They will be reconstructed and thus "heard" as a consequence of the remaining tonal movement of the hierarchically highest feature drawing upon the knowledge of the rules of tonal variation in Swedish with reference to the syllable structure of the utterance and thus, finally, to its duration.

DISCUSSION

Before outlining a new concept of a prosody model, some aspects of the interrelationships between time and Fo will be discussed.

Interrelationships between time and Fo

The results of the present investigation show that there is no simple relationship between the temporal and tonal structure of an utterance. Instead both structures are connected and interlocked with each other in different ways. Segment durations and Fo-movements that are to be found in the speech signal result partly from autonomous prosodic features of time and tone, partly from their mutual effects (tonal-to-temporal and temporal-to-tonal), and partly from intervening factors like speech tempo and the individual behaviour of the speaker. Thus, in conclusion, the relationships between duration and Fo in speech are not simple and unidirectional, but rather complex and bidirectional.

After all, is it really justified to ask the categorical question as to the independence of time and Fo-structure? A question which leads, for instance, to the view of the primacy of Fo. As a matter of fact, the tonal gestures or features, alone or in combination, are assigned to certain linguistic units like vowel, syllable, stress group, phrase, sentence, and text. Therefore the prosodic gestures are associated phonologically with linguistic units in some arrangement which is to be thought of as linear and punctual. However, even abstract tonal gestures are related to time because these linguistic units have to be projected onto the time dimension when manifested.

Time remains an autonomous dimension of prosodic features, even if we, as the gesture theory (Öhman et al. 1979, Engstrand 1983), assume that all the phonological gestures, spectral and prosodic, are not defined in temporal terms, but appear, for a given utterance, in an abstract string of simple or combined gestures and which are coarticulated together. When all these timeless gestures are executed, all of them, nevertheless, cannot come out as a natural consequence of their essential conditions and requirements. Some phonological gestures are, sui generis, temporal by nature. One and the same gesture may result in very different segment durations, according to context. The s-gesture, for instance, will be executed temporally in different ways depending on the segmental and prosodic context. Second, segment duration also varies in passages, context being constant, where Fo does not change and thus the constant Fo does not put any requirements at all on a given spectral gesture. Third, the view is generally accepted that the length distinction of quantity in Swedish and other languages, such as Danish and German, is tied to stress, i.e. quantity can only appear in stressed syllables. In this respect, a coarticulation of prosodic features exists. In unstressed positions, however, a reduction or neutralisation of quantity takes place, as is often the case with spectral gestures, for instance vowel reduction in English and Russian. In other quantity languages, like for instance Finnish and Czech, the quantity gestures also appear in unstressed syllables. Thus they are independent of stress in these languages.

The autonomy of temporal patterns will be even more obvious in a contrastive perspective. Take for instance the gesture or gestures for a voiceless, word-medial /p/ which we assume to be identical in languages, such as Finnish, Standard Swedish, Spanish, and Greek. It is a matter of fact that usually the /p/ in Greek and Spanish, absolutely and relatively, shows a much shorter duration than the /p/ in Standard Swedish or Finnish. In order to execute the essential element or elements of the p-gesture, a certain minimum of time is required. However, it is quite evident that the p-occlusion in Finnish and Standard Swedish is held longer than necessary, especially following a stressed short vowel in Standard Swedish or in a long consonant in Finnish, in order to be able to produce a good and complete /p/. Therefore everybody will realize that in certain languages, like for instance Swedish and Finnish, temporal gestures or features (quantity) are superimposed on spectral and tonal features in certain positions or, to put it in terms of the gesture theory, spectral and tonal features are coarticulated with temporal gestures, the latter, though, providing the basic and controlling frame of reference.

From a general linguistic view, it can be assumed that, in principle, temporal and tonal phonological features are autonomous in every language. The kind and degree of mutual relationships, of course, may vary from language to language. For instance, the tonal feature of sentence accent or emphasis does not increase segment duration in Danish (Thorsen 1980), in German it may do it optionally (Bannert 1982b), whereas in Swedish it will do so obligatorily. Danish does not show the decrease of vowel duration as a consequence of the increasing number of unstressed syllables in the stress group (Fischer-Jørgensen 1982) which is well witnessed in many languages.

Another problem with the model of the primacy of Fo (Lyberg 1981) where the duration of the stressed vowel is calculated from the change of Fo over this segment is the fact that, as a consequence, no segment duration can be calculated when there is no change of Fo over a given segment. The Fo-declination throughout the utterance is not considered as an Fo-change in this respect. A non-changing Fo is to be found in cases where Fo-points of the same level are concatenated low or high. Take, for instance, long compounds in Standard Swedish like <u>bòstadsbyggnadsprogramkommitté</u> where the low end of the word accent fall is connected low with the beginning of the rise of the phrase accent in the penultima or sentences with several syllables between accentuated ones: <u>Det var ju Pér som skulle ha skrivit brévet.</u> Here the tonal concatenation is high between the high tone of the phrase accent in <u>Pér</u> and the fall of the word accent in <u>bré</u>-. If, then, vowel duration cannot be calculated in such cases, let alone the duration of consonants, durations, in a model of speech processing, have to be taken from somewhere in order to assign typical and correct durations to all these segments. Even this consideration points to a solution treating durations and tonal movements as autonomous units.

Apart from the autonomous temporal and tonal features that, alone or in coarticulation, lay out their basic patterns in the time and frequency dimensions, one can observe various mutual effects of the prosodic features on each other at the phonetic level. These effects are temporal-to-tonal and tonal-to-temporal as well. A rising tonal movement usually takes more time than a falling one (cf. Ohala and Ewan 1973, Sundberg 1979; Elert 1964). This effect, however, is rather small compared to the total segment duration.

One effect of duration on Fo is due to speech tempo. Increased speech tempo leads to shorter segment durations and therefore there is less time available to execute the tonal gestures in addition to the spectral and temporal ones. Increased speech tempo, in general, increases Fo globally throughout the utterance and the range of Fo-variation is decreased (cf. Gårding 1975).

Outline of a model for prosody

The evaluation of the results of this study and the data and conclusions of other investigations (e.g. Thorsen 1980, Bruce 1981, Bannert 1982b) leads to the reasonable view that the temporal and tonal structures of an utterance, in one sense, are totally independent of each other. In another sense, however, they are connected and interlocked with each other. The temporal structure is considered primary, thus representing the necessary requirement for the execution or coarticulation of tonal features. Therefore, in a generative model of prosody, both dimensions, time and frequencey, have to be treated as autonomous dimensions, although mutual dependencies are to be found. What is important for the design of a prosody model is realizing that there are some separate temporal and tonal phonological features, thus that neither is derived from the other. The essential parts of the tonal contour of an utterance, expressed as tonal points or tonal movements, however, are projected onto the temporal structure which has been processed without any a priori dependence on tonal features.

Every prosody model represents only one part of a comprehensive speech model. The generation of prosody must not be seen as the last step in the derivation of the speech signal. Given the interlocking of the prosodic features with different linguistic components (pragmatic, semantic, syntactic, morphological, and phonological), it becomes quite obvious that the prosodic features have to be processed in a fully integrated way on several levels (cf. van Wijk and Kempen 1985). In accordance with this view of integration, the treatment of temporal and tonal structure isolated from other features and processes is abandoned and a comprehensive and interactive prosody model will be outlined. It is illustrated in Fig. 4.

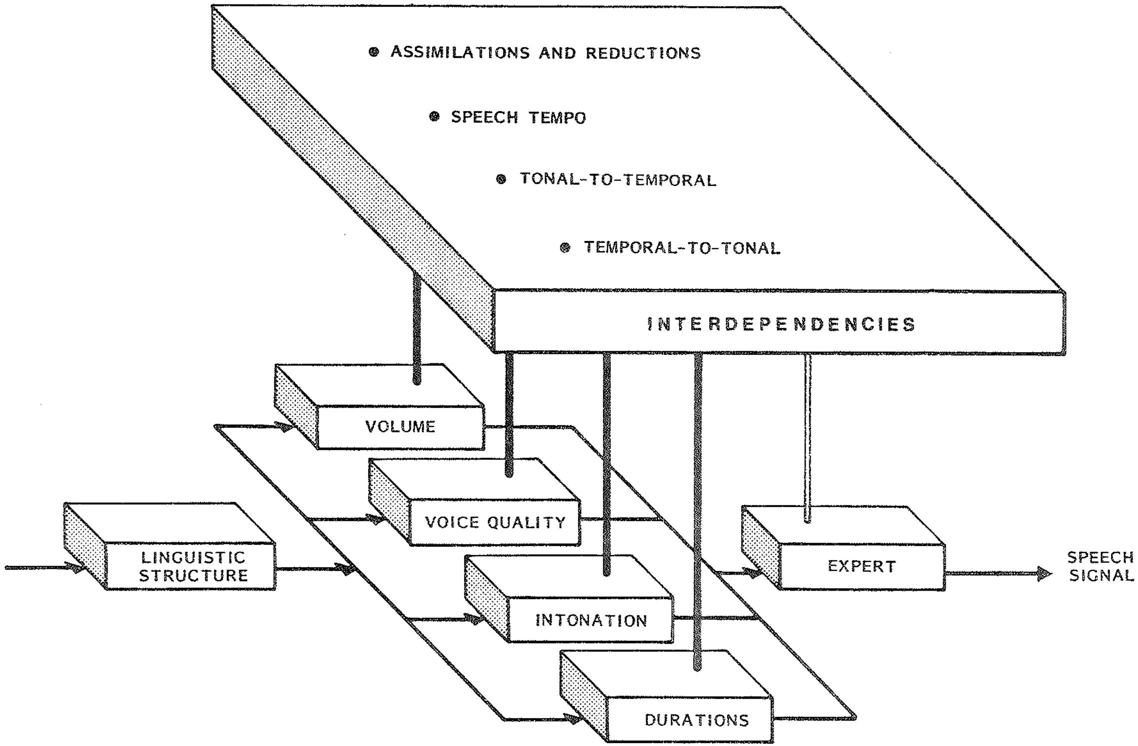


Fig. 4 Outline of a model for prosody where the prosodic structure of utterances is processed in parallel taking into consideration the different factors and relationships affecting the temporal and tonal structures of the output.

As before, the input of the prosody model is a linguistically fully specified string of linguistic units in a phonologically canonical form. The string is completely defined and contains all the necessary phonological (spectral, temporal, and tonal) features including voice quality and volume <5>, as well as the morphological, syntactic, semantic, and pragmatic features. In contrast to some prosody models, it is assumed that all relevant linguistic rules have operated before. Therefore I presuppose that all accent deletions, the assignment of sentence accent, etc. have already been done.

The processing of the information of the input is not done step by step where the output of one step serves as the only input to the next step. In a previous version of a prosody model, the basic temporal and tonal structures of an utterance were processed in a stepwise way, then added and finally, accounting for the mutual effects, modified in the modification component.

The design of the new model is based on the clear distinction between linguistic rules, information, and knowledge of various kinds. All linguistic rules and information including dependencies between features on different linguistic levels which are necessary for generating the prosodic structure of an utterance are available for the processing of the utterance simultaneously and continuously. Obeying the principles of applicability and utility, the rules and information are recalled and used whenever necessary and suitable. Fig. 4 shows the outline of a prosody model designed according to the principle of continuously flowing and complete information processing. Although the basic temporal and tonal structures, and the dimensions of volume and voice quality as well, are generated in separate channels, these sub-processes are effected and controlled all the time by the rhythmical and tonal rules, the information about the mutual effects of time and frequency, about speech tempo, assimilations and reductions (spectral arrangements). As a consequence of this, no further or final modification which otherwise would be necessary after the addition of the basic structure is needed. In the present version of the model, the prosodic structures are generated and processed in accordance with context and all the other relevant factors from the very beginning. It is immediately clear, however, that generating prosody in this on-line model amounts to a very complex process indeed. Nevertheless, the complexity of the processing of the speech signal should not be a deterrent argument against such an approach. On the contrary, it has to be assumed that the present outline of a prosody model is psychologically more realistic than the more simple and linear step-by-step model. In any case, the present design of the prosody model appears to be in rather good accordance with the essential ideas of coarticulation of gestures suggested in the gesture theory of Öhman et al. (1979).

It seems superfluous to remark, of course, that the present outline of a prosody model needs elaborating, completion and testing. However, the model in its present form makes it easier to formulate relevant and interesting questions. It also represents a new test programme in order to, in a coherent and dynamic model, investigate prosodic rules and interrelationships between different features in the time and frequency dimensions. Using speech synthesis by rule, it will be possible to optimize prosodic research by way of direct feed-back. The present model of continuous information processing for the generation of prosody can be connected with the representation of knowledge of artificial intelligence and with expert systems.

FOOTNOTES

<*> This research was supported by the Bank of Sweden Tercentenary Foundation.

<1> These are the high and low points which, in intonation models, are inserted as supporting points by rule with reference to segments or syllables , thus generating the Fo-contour step by step.

<2> For valuable help and discussion I am grateful to Klaus-Jürgen Engelberg, Olle Engstrand, Lennart Nordstrand, Gerhard Rigoll, and Herbert Tropf.

<3> No statistical testing was carried out for several reasons: The number of observations is too small and they could not be collapsed over variables and speakers.

<4> For the purpose of this comparison, it is immaterial that the first segment in position 1 is [an] of kan, whereas the first segment in positions 2 and 3 is [a] of lämna and långa, respectively.

<5> The dimensions of volume (intensity) and voice quality are included and indicated to complete the picture.

REFERENCES

Alstermark M. and Eriksson Y. 1972. Fundamental frequency correlates of the grave word accent in Swedish: The effect of vowel duration. STL-QPSR 2/3, 53-60

Bannert R. 1979. The effect of sentence accent on quantity. Proceedings of the 9th International Congress of Phonetic Sciences, Vol.II, 253-259. Copenhagen

Bannert R. 1982 (a). An Fo-dependent model for segment duration? Uppsala University, Reports from the Institute of Linguistics 8, 59-80

Bannert R. 1982 (b). Temporal and tonal control in German. Lund University, Department of General Linguistics and Phonetics, Working Papers 22, 1-26

Bannert R. 1983. Some phonetic characteristics of a model for German prosody. Lund University, Department of General Linguistics and Phonetics, Working Papers 25, 1-34

Bannert R. 1984. Towards a model for German prosody. Lund University, Department of General Linguistics and Phonetics, Working Papers 27, 1-36. Also: Folia Linguistica XIX, 321-341

Bannert R. and Bredvad-Jensen A-C. 1975. Temporal organization of Swedish tonal accents: The effect of vowel duration. Lund University, Department of General Linguistics and Phonetics, Working Papers 10, 1-36 and Working Papers 15, 133-138

Bannert R. and Bruce G. 1976. Svenska rytmer i ljud och bild. Umeå University, Department of Phonetics. Publication 10, 2-6

Bolinger D.L. 1958. A Theory of Pitch Accent in English. Word 14, 109-149

Bruce G. 1977. Swedish word accents in sentence perspective. Travaux de l'Institut de Linguistique de Lund XII. Lund

Bruce G. 1981. Tonal and temporal interplay. Lund University, Department of General Linguistics and Phonetics, Working Papers 21, 49-60. Also: Nordic Prosody II, 63-74. Fretheim T. (ed.). Trondheim

Bruce G. 1983. On Rhythmic Alternation. Lund University, Department of Linguistics and Phonetics, Working Papers 25, 35-52

Elert C-C. 1964. Phonological studies of quantity in Swedish. Uppsala

Engstrand O. 1984. Articulatory Coordination in Selected VCV utterances: A Means-End View. Uppsala University, Department of Linguistics. Report 10

Engstrand O., Nordstrand L., and Zetterlund S. 1978. Some observations on the role of prosodic parameters in the perception of phrase structure in Swedish. Phonetics Symposium, 11-13. Nord L. and Karlsson I. (eds). Speech Transmission Laboratory, Royal Institute of Technology, Stockholm

Fischer-Jørgensen E. 1982. Segment Duration in Danish Words in Dependency on Higher Level Phonological Units. University of Copenhagen, Annual Report of the Institute of Phonetics 16, 137-189

Gårding E. 1975. The Influence of Tempo on Rhythmic and Tonal Patterns in Three Swedish Dialects. Lund University, Department of Linguistics and Phonetics. Working Papers 12, 70-83

Gårding E. 1981. Contrastive Prosody: A Model and its Application . Studia Linguistica 35, 146-165

Gårding E., Botinis A., and Touati P. 1982. A Comparative Study of Swedish, Greek, and French Intonation. Lund University, Department of Linguistics and Phonetics. Working Papers 22, 137-152

Gårding E. and Bruce G. 1981. A presentation of the Lund model for Swedish intonation. Nordic Prosody II, 33-39. Fretheim T. (ed.). Trondheim. Also: Lund University, Department of General Linguistics and Phonetics, Working Papers 21, 69-75

Gårding E., Bannert R., Bredvad-Jensen A-C., Bruce G.,and Nauclér K. 1974. Talar skåningarna svenska? Svenskans beskrivning 8, 107-117. C.Platzack (ed.). Lund

Klatt D.H. 1975. Vowel lengthening is syntactically determined in a connected discourse. J. of Phonetics 3, 129-140

Lindblom B. 1979. Final lengthening in speech and music. Nordic Prosody, 85-101. Gårding E., Bruce G., and Bannert R. (eds). Travaux de l'Institut de Linguistique de Lund XIII

Lindblom B., Lyberg B., and Holmgren K. 1981. Durational patterns of Swedish phonology: Do they reflect short-term memory processes? Indiana University, Linguistics Club

Lyberg B. 1981. Temporal properties of spoken Swedish. Monographs from the Institute of Linguistics 6, Stockholm Univesity

Öhman S., Zetterlund S., Nordstrand L., and Engstrand O. 1979. Predicting segment durations in terms of a gesture theory of speech production. Proceedings of the 9th International Congress of Phonetic Sciences, Vol.II, 305-311. Copenhagen

Ohala J. and Ewan W. 1973. Speed of Pitch Change. JASA 53, 345 (A)

Sundberg J. 1979. Maximum Speed of Pitch Changes in Singers and Untrained Subjects. Journal of Phonetics 7, 71-79

Svensson S-G. 1974. Prosody and Grammar in Speech Perception. University of Stockholm, Monographs from the Institute of Linguistics 2

Thorsen N. 1980. Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish. University of Copenhagen, Institute of Phonetics, Annual Report 14, 121-205

van Wijk C. and Kempen G. 1985. From sentence structure to intonation contour. An algorithm for computing pitch contours on the basis of sentence accents and syntactic structure. Germanistische Linguistik 79-80, 155-182. Sprachsynthese, Müller B.S. (Hrsg.). Hildesheim

Zetterlund S., Nordstrand L., and Engstrand O. 1978. An Experiment on the Perceptual Evaluation of Prosodic Parameters for Phrase Structure Decision in Swedish. Nordic Prosody 15-22. Travaux de l'Institut de Linguistique de Lund XIII. Gårding E., Bruce G., and Bannert R. (eds). Lund

The prosodic structure of Greek

A phonological, acoustic, physiological and perceptual study

Antonis Botinis

Summary and results

0 Outline

This study deals with Greek prosody and is composed of four parts. In the first part (1) the phonological system of Greek prosody at the lexical level, word level, phrase level and sentence level is described and the rules to assign lexical stress, word stress, phrase stress and sentence stress to the corresponding level are presented.

The second part (2) is a report of three acoustic experiments. The purpose of experiment I was to investigate the contribution of the three acoustic parameters of fundamental frequency (Fo), duration and intensity to the construction of word and sentence stress. Experiment II was to investigate the three parameters' contribution to phrase stress and compare the acoustic manifestation of an enclitic structure to one of a proclitic structure. Experiment III was to investigate two different syntactic structures' acoustic manifestations and examine their relation to prosody.

The third part (3) is a report of a physiological investigation of the variations of subglottal air pressure (Ps) associated with word stress in post-focal position. The purpose of this experiment was to find out how the acoustic parameters co-vary with subglottal pressure and if subglottal pressure affects one or more acoustic parameters, to what degree.

The fourth part (4) is a report of six perceptual experiments. The purpose of experiment I was to find out which of the acoustic parameters contributes most to the perception of word stress after focus. Experiment II was to find out which of the acoustic parameters contributes most to word stress perception before focus. Experiment III was designed to test if Fo, which contributes most to pre-focal word stress, is perceived categorically or continuously in the vertical dimension; experiment IIII was to test if Fo is perceived categorically or continuously in the horizontal dimension. Experiment V was to find out which of the acoustic parameters contributes most to the perception of phrase stress, and experiment VI was designed to test if Fo is perceived categorically or continuously.

1 Greek Prosodic Phonology

In this study, language is thought of as a complex entity of different levels of representation. Each level is associated with its corresponding prosodic category, i.e., the lexical level with lexical stress, the morphological level with word stress, the syntactic level with phrase stress, and the semantic and textual level with sentence stress. The lexical stress is given by the lexicon and may appear on any one of the last three syllables. The word stress may coincide with the lexical stress although it usually moves to the right whenever the word boundary moves to the right by the addition of extra syllables through inflection and derivation or when some morphemes attracting word stress are added to the lexical word. The phrase stress appears two syllables to the right of word stress when the phrase boundary is more than two syllables to the right of word stress. The sentence stress appears on the last lexical element bearing focus. However, phrase stress attracts sentence stress within the domain of the phrase no matter which element is in focus.

The prosodic categories are organized into a hierarchical structure and have a classificatory function, either they exist or not, and every category has its own distribution rules which apply to the corresponding level. The freedom of the prosodic categories varies according to the level on which their rules are applied. The lower the level the less dependent the prosodic rules are upon the higher levels of representetion. In other words, word stress needs information only about the position of lexical stress and the word boundary whereas phrase stress, apart from word stress information, needs information about phrase boundaries as well, i.e., a higher level is involved. Thus, prosody cannot be independent from the morpholexical, syntactic and semantic structure of the language; it is rather an abstract linguistic entity which the different prosodic categories create with concrete contributions from the corresponding levels of representation.

2 Acoustic Analysis

In acoustic experiment I the parameters of Fo, duration and intensity were found to contribute to pre-focal word stress, only duration and intensity to post-focal word stress and Fo, duration and intensity to sentence stress. Of the three acoustic parameters examined, duration and intensity run parallel to each other and are usually present for both word and sentence stress. The word stress acoustic parameters are not constant across the syntagm but their manifestation is organized in relation to the sentence stress position.

In experiment II phrase stress is manifested by the rising of Fo at the pre- and focal position and the falling and flattening of Fo at the post-focal position combined with longer duration and higher peak intensity. At non-focal position the acoustic structure of phrase stress is the same as the word stress, the difference between the two categories being of perceptual as well as of functional nature. Fo is not a strong enough perceptual cue for the phrase stress distinction the way it is for the word stress; on the other hand, word stress operates at the morpholexical level to distinguish contrastive words whereas phrase stress operates at the syntactic level to distinguish contrastive structures. Apart from phrase stress, the enclitic structure appears with the same acoustic manifestation as the proclitic one.

In experiment III the influence that syntax may have on prosody has been corroborated. Two noun phrases with the same number of syllables, in the same context, one with a word and a phrase stress and the other with two word stresses may apparently have the same acoustic manifestation outside focus; but when focus is involved the noun phrases take completely different acoustic manifestations. The application of sentence stress on the lexical entity with a word stress is blocked when the lexical entity is in a phrase where there exists a phrase stress, whereas sentence stress may be applied to any one of the lexical entities composing the phrase and have a word stress.

3 Physiological Analysis

In the physiological experiment the Ps was found to co-vary with intensity. It seems, then, that the larynx is mainly responsible for Fo variations and the subglottal system for intensity. Intensity was found to correlate with Fo for sentence stress but with duration for post-focal word stress, an implication that the acoustic parameters are independent of each other and not produced by the same mechanism.

4 Perceptual Analysis

In the perceptual experiment I duration was found to be the most important acoustic cue although a combination with intensity was necessary for the perception of word stress at post-focal position. Listeners could perceive word stress distinctions after focus where the acoustic parameters are weakly manifested even with synthetic speech. The context, Fo, and formant structure did not have any perceptual contribution at this position.

In experiment II Fo was found to be the all important perceptual cue for the pre-focal word stress distinction. A hierarchy of the acoustic parameters is still questionable since Fo is the primary cue and overrides duration and intensity as conflicting cues.

In experiment III the question of whether Fo is perceived categorically or continuously in the vertical dimension for the pre-focal word stress distinction is still unanswered. However, the main finding of this experiment was that the prosodically contrastive words kept their original meaning even with a neutralized Fo-contour. It seems as if Fo functions as a distinctive feature with intensity and duration as redundant features and when the distinctive features are neutralized, the redundant ones take over. Thus, Fo is not an absolute necessity for word stress perception; duration and intensity may equally effectively convey the word stress concept at pre-focal position the same way they do at post-focal position.

In experiment IIII in which Fo had been moved horizontally across the prosodic minimal pair under investigation at pre-focal position, word stress was perceived categorically. Fo had to be far away from the mid-point of the syllable to neutralize the influence of duration and intensity from the original word. This finding reflects the influence that duration and intensity may have on pre-focal word stress perception.

In experiment V listeners could perceive phrase stress and associate it with the proper syntactic and semantic structure. The phrase stress Fo-rise contributes most to phrase stress perception, but it has to be combined with duration and intensity to denote the concept of phrase stress in contrast to word stress in pre-focal position where Fo itself is the decisive factor. In the light of these findings it seems that the idea that a certain parameter does not contribute to stress perception should be reconsidered. A particular parameter may not contribute by itself at all; however it may be decisive when combined with the other acoustic parameters.

In experiment VI the question of whether Fo is perceived categorically or continuously for the phrase stress distinction is still unanswered. However, the present experiment corroborates the results of experiment V where Fo was not enough to convey phrase stress by itself, and experiment III where intensity and duration became distinctive features when Fo was neutralized.

Notes on some particle and prepositional constructions in Swedish and English

Helen Goodluck

This paper concerns mainly the constituent analysis for particle verb constructions. In sections A-D, a complex verb analysis for particle verb constructions in Swedish is proposed, and particle constructions are distinguished from prepositional constructions with the same stress pattern. In section E, a complex verb analysis for particle constructions in English is defended against arguments by Kayne 1985 against such an analysis. A difference in the distribution of double object constructions in the two languages can be made to follow from the existence of a particle movement rule in English.

A. THREE SWEDISH SENTENCE TYPES

It is widely accepted that the 'particle' in particle verb constructions in English and Swedish should be analyzed as a member of the category preposition, or possibly adverb in some instances (Emonds 1972; Ejerhed 1979, n.d.). In what follows we will use the term 'preposition' to refer to a preposition that is head of a PP and the term 'particle' for a preposition that is part of a complex verb construction. The following three sentence types will be distinguished in Swedish:

i. Particle verb constructions, where the meaning of the verb plus particle is non-compositional in many cases and where there is compound stress (destressing to the left), resulting in primary stress on the particle,[1]

1. Flickan tǒg áv plåstret (´= stress degree n
   'the girl took off the bandaid' ˇ= stress degree <n)

ii. Prepositional constructions that have a stress pattern similar or identical to that of the particle constructions and a fairly predictable semantic interpretation. In

2. Katten hǒppade bákom elefanten
   'the cat jumped behind the elephant'

the object NP is interpreted as the location towards which the cat jumps (i.e. the interpretation is 'jump to behind').

iii. Prepositional constructions without stress on the preposition, where the object of the preposition is interpreted as the location at which the action takes place. The sentence

3. Katten hóppade băkom elefanten

is interpreted to mean that the cat jumped up and down behind the elephant.[2]

B. SYNTAX

The operation of gapping provides evidence that sentences such as (2) are to be grouped with sentences such as (3), rather than sentences such as (1), despite the shared stress pattern of (1) and (2). The paradigms for gapping of the verb vs. verb plus particle/preposition for sentences of the types (1), (2) and (3) are given in (4), (5) and (6), respectively. As the judgements there show, the P must gap with the verb in sentences of type (1) and may not gap with the verb in sentence types (2) or (3).

4.a Per tog av sin hatt och Kalle tog av sin halsduk
'P. took off his hat and K. took off his scarf'
b Per tog av sin hatt och Kalle Ø Ø sin halsduk
c *Per tog av sin hatt och Kalle Ø av sin halsduk[3]

5.a Katten hŏppade bákom elefanten och hunden hŏppade bákom kon
'The cat jumped behind the elephant and the dog jumped behind the cow'
b *Katten hŏppade bákom elefanten och hunden Ø Ø kon
c ?Katten hŏppade bákom elefanten och hunden Ø bákom kon

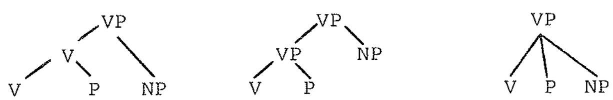
6.a Katten hóppade băkom elefanten och hunden hóppade băkom kon
b *Katten hóppade băkom elefanten och hunden Ø Ø kon
c Katten hóppade băkom elefanten och hunden Ø băkom kon

Some, but not all speakers, accept gapped sentences of type (2) where the verb only is removed in the second sentence and the preposition is stressed (and assign the object in the gapped sentence the towards-location interpretation); the question mark on (5c) is returned to below.

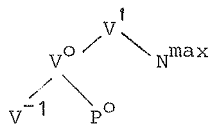
From the gapping test it can be inferred that the P in sentences of type (2) is head of a PP, as it is in (3). Less can be said about the exact constitutent structure of (1).

In Stillings' (1975) analysis based on English, gapping is constrained to delete a sequence of one or more identical verbs in the second conjunct, requiring also that the material to the right of the gap site(s) be a single constitutent. On this analysis, we can infer both that the P--NP sequence to the right of the gap site in (5c) and (6c) is a PP (single constituent) and that the V--P sequence that is gapped in (4b) is a complex verb $[_V$tog[av]], since only members of the category verb gap in Stillings' analysis. (The relevance of gapping to complex verb status on Stillings' analysis is pointed out by Selkirk 1982, p.28). However, the facts are not more than consistent with a complex verb structure. Stillings herself must admit cases of 'reanalysis', where, for example, an NP is reanalyzed as part of the V (to allow 'John writes poems in the bathroom and Sue ~~writes poems~~ in the garden', where the material struck through is the sequence to be gapped; parallel examples exist in Swedish). Moreover, the facts in (4-6) could all be accounted for by a tighter restriction on the material to the right of the gap site, to the effect that it is not only a constitutent, but an immediate constituent of a projection of V (cf. Neijt 1979, Cht 3); that will exclude deleting the head of a PP (5b/6b) and will permit a direct object to be left (4b), without any inference about the syntax of the material that is deleted.[4]

We have then three logical possibilities for the syntax of the V--P--N sequence in (1): a complex verb structre, a hierarchical structure where P is attached to a phrasal projection of V at a lower level than the object, and a flat structure analysis,



Ejerhed (1979,n.d.) opts for the third, flat structure, analsis, as opposed to a complex V analysis, on the ground that a complex V analysis assumes rules that incorrectly predict recursion of the internal V, with complexes of P on the right; such particle constructions do not occur.[5] She notes also that this potential problem would not arise if some other symbol than plain V were expanded by the rule for forming the complex verb. Recent studies of morphology suggest that such a restriction may be accommodated within general constraints on word-formation. Thus, following Selkirk's (1982) application of X-bar theory to word-structure, we can propose that verb particle constructions are the result of a rule where the category V at bar level 0 (the level WORD) is expanded to V at level -1 (the level STEM) and P at 0. The structure for the verb phrase is (1) will thus more precisely be,



and the possibility of recursion will be excluded. This analysis is consistent with Selkirk's general constraint on word structure rules (p. 8) that a word category can only be rewritten with categories at the same or lower bar level, but is at odds with Selkirk's (tentative) analysis of verb-particle constructions as compounds, which on her theory must be composed of categories that bear the same bar level specification as the dominating category (see pp. 47-52). However, the analysis does away with the recursion problem (a problem that Selkirk fails to note for her treatment of English verb particle constructions as compounds); moreover, the removal of particle verb constructions from the inventory of compound types eliminates the one counterexample in Selkirk's analysis to the generalization that English compounds are right-headed.

It is difficult to find any clear cut ground for choosing between the two hierarchical versions of the structure for particle constructions (complex verb and phrasal projection of V). On the side of the complex verb analysis (and against both the hierarchical and flat structure phrasal analysis), the complex verb structure is consistent with a theory of X-bar word and phrase syntax where all complement positions at the phrasal level are maximal projections, with the expectation that they will show the full range of complementation for the category type (i.e. it eliminates the need for a constraint to ensure only intransitive prepositions occur between the V and object, a problem Selkirk (p. 28) notes for the flat structure analysis of particle verbs in English). A potential problem for the complex V analysis in Swedish is that the particle does not move with the verb (e.g. in subject-verb inversion, 'Tog flickan av plåstret?', *'Tog av flickan plåstret?'), this can be handled at least mechanically in the X-bar analysis above by requiring rules that move V to affect Vmin (or the finite verb, leaving aside any problems with feature percolation). In Section E assuming some kind of hierarchical analysis in the base for both Swedish and English will contribute to an account of a difference in distribution of double object constructions in the two languages. We will simply adopt the complex V analysis as a working hypothesis.

It might be argued that a flat structure analysis is needed, and is the correct analysis for sentences of type (2). That analysis would account for the fact that some speakers reject sentences such as (5c), on the assumption that the material to the right of the gap site must be a constituent. Additionally, when the preposition and object are preposed to initial position, the interpretation given to the object is almost invariably that of 'in location' (the reading corresponding to (3)), with stress on the preposition in the preposed phrase being rejected or interpreted as contrastive stress of some type.

7. Bakom elefanten hoppade katten.

The absence of preposition stressing and the towards-location reading would follow on the flat analysis, given the standard assumption that only constituents can be preposed.

However, I do not think a flat structure analysis is necessary for sentences of type (2). With respect to gapping, the rejection by some speakers of sentences such as (5c) may reflect an additional constraint in their dialects, that requires the eliminated materials to be a complete semantic unit. Such a constraint cannot be met in (5c) consistent with the syntactic conditions on gapping (whichever version of the conditions sketched above is adopted).

With respect to the interpretation of preposed PPs, it is worth noting that constituenthood is not a necessary condition for unifications by stress, which is the cue for the toward location reading for sentences such as (2).[6] But linear continguity may be. A simple solution to the absence of the 'toward location' reading for sentences such as (7) would be that the prosodic pattern of stress on the preposition is assigned at level 'after preposing'. However, given the evidence that the results of such stressing operations are frequently preserved in the outputs of reordering (see Rischel 1983 and references therein), something more sophisticated (or at least different) will need to be said about why the towards location reading is absent for (7) for most speakers. One possibility is that the stylistic function of preposing gives priority to a contrastive interpretation of stress on the preposition, and thus indirectly promotes the in-location reading.[7]

C. SEMANTICS

Ejerhed (n.d.) describes several regularities characterizing the meaning of verb particle constructions that do not have lexicalized (completely non-compositional) meaning. The particle may perfectivize and may transitivize the verb to which it is added; it may also imperfectivize and intransitivize and it may add various completative nuances of meaning. The following are among the examples given in her paper (p. 21-22),

8.a Perfectivization

Vattnet rann	'the water was running'
Vattnet rann ut	'the water ran off'

b Transitivization and perfectivization

Hon satt	'she was sitting'
Hon satt av föreläsningen	'she sat through the lecture'

c Intransitivization

Han såg matchen	'he saw the fight'
Han såg på	'he watched'

d Completive meanings

Han åt kakan	'he ate the cake'
Han åt upp kakan	'he finished the cake'
Huset brann	'the house was on fire'
Huset brann upp	'the house burned down'
Han sköt två soldater	'he shot two soldiers (accidentally or intentionally)
Han sköt ned två soldater	'he hit and killed two soldiers'

We can add the observation that where the meaning of the particle remains fairly transparent, the interpretation is frequently causative or resultative. Thus in 'Jan tog på hatten' (Jan put on the hat), the hat is on as a result of the action.

The alternation in meaning between the prepositional examples in (2) and (3) is different from any of these functions, and can be expressed in terms of a change in the thematic role assigned to the object NP. Following Gruber (1976) and others, we can designate the role of the object in the towards location interpretation in (2) as GOAL and the role of the object in the in-location interpretation of (3) as LOCATION. We assume that individual prepositions that permit the stress alternation are lexically specified for either a GOAL or LOC object (cf. the English prepositional equivalents, such as the gloss for (2), which are ambiguous between the two readings); preposition stressing in Swedish must then be associated for this construction with a semantic rule that selects GOAL as the thematic role of the object. Intuitively, this association of goal (rather than LOC) with the stressed preposition is non-arbitrary, in that the stress unifies the verb and the preposition, and LOC is not (canonically) assigned to non-prepositional arguments in Swedish; but since I have argued in section B that the object remains syntactically a prepositional object in sentences such as (2), such a generalization cannot be cashed out at the level of surface syntax.

D. SUMMARY: COMPOUNDING VS. MODIFICATION IN WORD STRUCTURE

On the analysis given above, Swedish V--P--NP sequences conform to current (X-bar) conceptions of word structure and phrase structure. At the word level, words are composed of categories at the same or lower bar-level than the category of the word itself ($V^{o} \rightarrow V^{-1} P^{o}$). At the phrasal level complements are maximal projections ($V^{n>o} \rightarrow \ldots P^{max}$). Unification by stress is not a sufficient condition for inferring particle verb syntax.

There are other facts concerning particle constructions to be examined in the context of the type of word-syntax used here, which may throw light on the relation between the structure of words and their semantic interpretation. For example, there is an alternation in some cases between a verb with particle in final vs. initial position. The initial position variant will tend to be less colloquial or to have a more abstract meaning. Thus Ejerhed (p. 23) gives the following examples:

transportera ut = uttransportera
'transport out'; synonymous, transporera ut more colloquial

bryta av ≠ avbryta
'break off' 'interrupt'

If we adopt the analysis above for verb plus particle sequences and a compound analysis for particle plus verb (the latter along the lines of Selkirk's analysis of English P--V verbs), the contrast will be one of difference vs. sameness of bar levels of the constituents of the verb,

Intuitively, this difference in syntax can be made to fit with the concrete-abstract contrast in meaning in the following way. In the V--P construction the P is a modifier of the head whereas in the P--V construction it is a subconstituent of equal rank as the head (where headedness is determined by shared category features and modifier defined as a constituent of the word with bar level distinct from the head). Modifiers may be expected to add to the meaning of the head, but not to change its basic meaning; by contrast, the meaning of compounds may be determined by rules in which each constituent of the compound contributes to the basic meaning, with the possibility of a shift away from a concrete meaning of the verbal head.

E. KAYNE'S ARGUMENTS AGAINST COMPLEX V IN ENGLISH

The observations made above with respect to the syntax and interpretation of Swedish particle constructions largely apply to English particle constructions also. English particle constructions show the same pattern of gapping,

9.a Sue took off the bandaid and Fred took off the cast
  b Sue took off the bandaid and Fred Ø Ø the cast

and similar semantic patterns in both preverbal and post verbal position. English differs from Swedish in permitting particles to follow as well as precede the direct object,

10.a Fred took off the cast
  b Fred took the cast off

and in permitting double object particle sentences with a predicative or dative interpretation, as in (11) and (12) respectively,

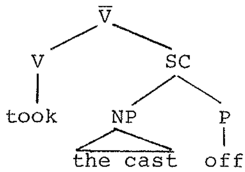
11. John made Bill out a liar

12. John handed Bill down the tools.[8]

Both the alternation between sentence pairs such as (10a, b) and the derivation of double NP particle sentences are dealt with by Kayne (1985), who analyses particle constructions in English as instances of small clause constructions, within a Government-Binding framework. Here I will outline Kayne's analysis and evaluate the arguments he makes against a complex verb analysis.

In Kayne's analysis, the D-structure for particle constructions has the particle in final position, as head of a small clause, of which the post-verbal NP is subject. Thus the D-structure of the VP for both 10a and 10b will be 13,

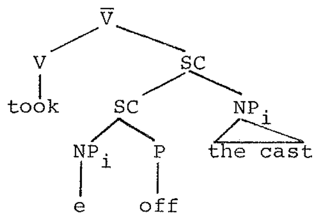
13.



(SC = small clause)

The S-structure for 10a will be derived by rightwards movement of NP, to yield 14,

14.

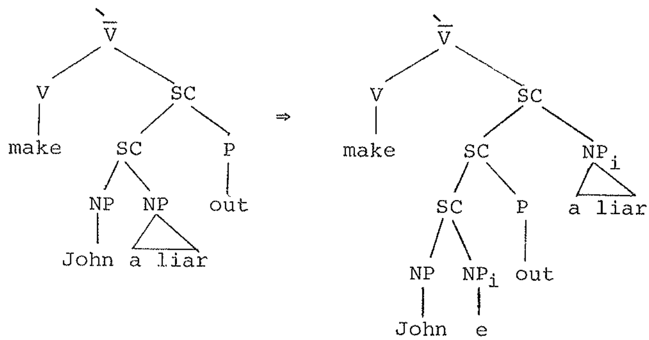


Kayne's analysis covers predicative and dative double NP constructions, in a manner that accounts for the fact that for many speakers the middle position is the only acceptable position for the particle (for all speakers it is the preferred position, cf. Emonds 1972, for discussion with respect to datives),

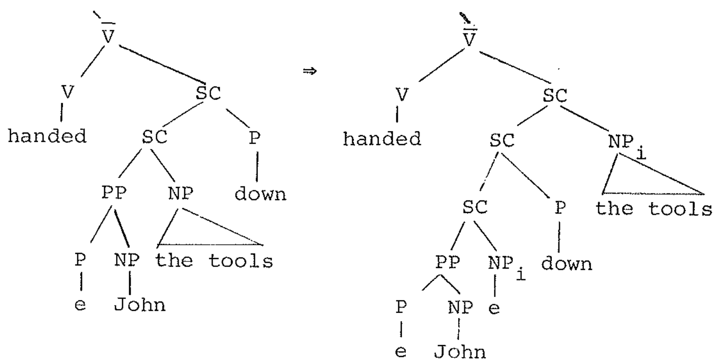
15.a John made Bill out a liar (=11)
b *?John made out Bill a liar
c *John made Bill a liar out

16.a John handed Bill down the tools (=12)
b *?John handed down Bill the tools
c *John handed Bill the tools down

In Kayne's analysis, both constructions involve a D-structure with a double small clause, as shown in the structures in (17) and (18). The S-structure order is derived by movement of NP, as in the case of (10b), 17.



18.



(P_e = empty preposition, source of theta-role for both NPs in Kayne's analysis).

Under Kayne's analysis, rightwards movement of the NP, as shown in (17) and (18), is necessary if the S-structures are to be such that case can be assigned to the NPs (the movement ensures that no more than one maximal projection containing lexical material intervenes between the verb and each NP, allowing case to be assigned in a way not possible in the unmoved structure; see Kayne 1985, sections 3.3 and 4.1 for details). In this manner, Kayne accounts for the fact that only the (a) version of (15) and (16) is grammatical for most speakers.

Kayne gives six points against a complex V analysis for particle verb constructions with the particle in post-verbal position; he argues that the facts that present potential problems for a complex verb analysis will follow from the alternative small-clause and NP movement analysis sketched above.

Kayne's arguments against a complex verb analysis are as follows:

[1] A complex verb analysis does not explain why inflectional morphemes cannot occur on the complex word,[9]

19. *John look up-ed the information 128

This will follow on Kayne's analysis, since the verb and particle are not a unit of category verb.

[2] A complex verb analysis does not account for why complex verbs cannot be followed by a pronominal object; thus (20) is ungrammatical on the reading where *it* is taken as a direct object of *look up*, rather than as the object of the preposition *up*,

20. *John looked up it. 129

In Kayne's analysis, (20) would be derived by rightwards movement of the pronoun from within a small clause headed by *up*; Kayne argues that the ungrammaticality of (20) is a reflex of a general filter on the output of rightwards movement, which requires that the moved NP be "heavier" than material it moves over (see Kayne, p. 127); particles will have a weighting higher than pronouns, and (20) will thus violate this filter.

[3] A complex verb analysis, Kayne argues, does not account for the fact that for a majority of speakers complex verbs have to be excluded from predicative small clause constructions or dative constructions, such as (15b) and (16b). Under the complex verb analysis and normal assumptions about subcategorization, we would on the face of it expect such constructions to be possible. As noted above, the ungrammaticality of such sentences, and the mandatory middle position for the particle (15/16a) follow on Kayne's analysis from conditions on the assignment of case.

[4] Similarly, Kayne argues that complex verbs would have to be excluded form cooccurring with infinitives having idiomatic subjects or *there* as subject. While the sentences in (21) are grammatical, those in (22) are ungrammatical,

21.a (?)They are trying to make out John to be a liar 59
b They are trying to make John out to be a liar 60

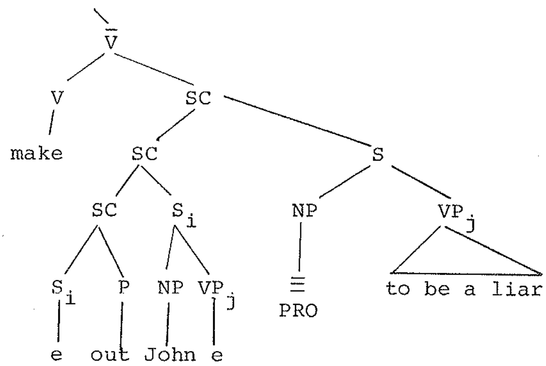
22.a *They are trying to make out advantage to have been taken of them 66
b *They are trying to make advantage out to have been taken of them 67

c *They are trying to make out there to be no solution to this problem 68

d *They are trying to make there out to be no solution to this problem 69

Kayne's analysis for (21-22) is as follows. Sentences such as those in (21a) will be derived by rightwards movement of the sentential subject of the small clause headed by <u>out</u>, followed by movement of the infinitival VP, with insertion of a subject that is effectively a PRO, creating a control structure. The S-structure of a sentence such as (21a) will thus be,

23.



Sentences such as (21b) can be derived with movement of the infinitival VP only. The sentences in (22) will be ruled out by a general prohibition on idiom chunks and <u>there</u> as controllers (cf. from Kayne, p. 115, *'There were reptiles before being mammals'; *'Advantage was taken of John's inattention before being taken of his stupidity')

[5] Kayne's fifth argument concerns wh-movement. With plain (non-particle) verbs, the second object in a double NP construction can be questioned,

24.a We handed John the tools.

b Which tools should we hand John first? 133

Kayne observes that even for speakers who accept sentences such as (16b) (=25a), (25b) is ungrammatical, in contrast to (24b),

25.a We handed down John the tools

b *Which tools should we hand down John first? 134

Kayne's argument is that if (24a) and (25a) have essentially the same structure, as the complex verb analysis claims, there should be no contrast between (24b) and (25b). On Kayne's analysis, the ungrammaticality of (25b) follows from movement of the wh-phrase from its derived structure position, adjoined to the small clause projection of *down*; (cf. the structure in (18) above). That position is an $\bar{A}$ position, which may not be a variable site.

[6] Kayne's final argument concerns the intensifier *right*, which can occur before P in particle constructions, but not before P in compounds,

26. ?John looked right up the information 135

27. *John right upended the rocking chair 136

The argument is that the distinction is to be expected if *right* cannot modify constituents of complex words, and V--P sequences are not complex verbs.

The following alternatives to Kayne's points can be proposed, under a complex verb analysis of particle constructions.[10]

Kayne's point [1] loses some of its force if we adopt the modificational structure for verb particle constructions sketched above (section D), where the head V is at a bar level lower than that of the particle that modifies it. Inflectional endings will be placed on the head of the word (V^{-1} for particle verbs), excluding *look up-ed*. The *right* argument (point [6]) might be handled in a similar manner. The correct distribution for *right* could be obtained if *right* is a specifier to words at category level n, and particles in particle verb constructions are at the requisite level and prepositions/particles in compounds are below it. However, this would require some change in the proposal above that particles in particle verb constructions and compounds are both constituents at level 0 (word), with a concomitant change in the assumption that constituents of words are always at bar level 0 or lower (if P level 1 is admitted into particle constructions).[11]

Kayne's argument [2], concerning the ungrammaticality of 'John looked up it', loses force in view of sentences such as (28),

28. *John gave Bill it.

In Kayne's analysis of datives (cf. the structures for (16a) above and Kayne 1984, Chts 7 and 9), it should be possible to base-generate such sentences, and their ungrammaticality cannot beattributed to a "weight" filter on the output of rightwards movement. Examples such as (28) suggest that the filter may apply more generally to block a pronoun object occurring to the right of a heavier constituent in a manner that potentially will cover both (28) and (22), with a complex verb analysis for the latter.

Kayne's third argument is that there is an absence of constructions such as (15b) and (16b) ('John made out Bill a liar'; 'John handed down Bill the tools'), contrary to expectations under a complex verb analysis and normal assumptions about subcategorization. In response to this argument, we can propose a different analysis, which will both permit particle verb structure and require the particle to follow the first object in such constructions. We will assume the base structure to be one in which at least the first NP is sister to V^o. We will take the dative case as the paradigm case, since there the meaning of the particle verb is maximally compositional, and make the following assumptions. In the internal structure of complex verbs, each constituent can separately specify the thematic role of an argument. For hand down, hand will specify THEME and down will specify GOAL. The following condition will apply: if an argument is syntactically realized, it must be c-commanded in S-structure by the element that determines its thematic role.[12] If the NP-P order in English particle constructions is derived from a complex V D-structure by virtue of a rule of particle movement that raises the P to a position under V1, then that rule will make legitimate double object constructions with verbs such as hand down. The contrast between (16a) and (16b) is thus accounted for by the c-command condition on theta-role assignment, plus the existence of a particle movement rule that raises the particle into a position where it c-commands the first object. This analysis predicts that if a language had complex (particle) verb constructions, but did not have a particle movement rule, then there would be no sentences such as (15b) and (16b), since there would be no way to make the S-structure conform to the c-command condition. Swedish appears to fit this prediction. It does not have particle movement, and, as far as I can tell from questioning my informants, it does not have double object constructions with particle verbs.[13]

The analysis of the middle placement of the particle in double NP constructions given here is intuitively most plausible where the relation between the particle and the NP to which it putatively assigns a thematic role is fairly transparent (so, for example, (16b) has the paraphrase 'John handed the tools down to Bill', where 'down to Bill' can have PP status ('It was down to Bill John handed the tools')). It can be counted as a mark in favor of this approach that the middle position preference may be relaxed where the meaning of the verb plus particle is relatively non-compositional, and it is less plausible that the particle assigns a thematic role independently of the verb. This relationship between how close-knit the verb and particle are and placement of the particle was noted by Bolinger (1971, p. 179), who finds the following sentences equally acceptable,

29.a Pack your brother up a nice lunch
b Pack up your brother a nice lunch.[14]

Kayne's fourth argument concerns idiom chunks and *there* in sentences such as those in (22). His account is that rightwards movement of the infinitive from within a small clause creates a control structure, which is illegitimate for sentences with idioms and *there*, since idiom chunks and *there* are in general barred from being controllers. An alternative account under a complex V analysis is to posit a *tell*-type control structure in the base for particle verb constructions with infinitival complements; the D-structure for (21a) and (22a) will then be,

30. They are trying to make out [John] [PRO to be a liar]

31. They are trying to make out [advantage] [PRO to have ...]

The sentences in (22) can then be ruled out on the same ground that Kayne excludes them (the bar on idiom chunks and *there* as controllers), but with an analysis that allows complex verbs.

Kayne notes that idioms and *there* are better in infinitival comlements to particle verbs when the particle construction is passive,

32. ?Advantage was made out to have been taken of them 72
(?)There was made out to be no solution to this problem

a fact that Kayne atrributes to passive having placed the idiom chunk/*there* in a position where it c-commands the subject of the infinitive, which can then be interpreted as a trace bound by the subject, rather than as PRO (cf. the structure 22). An alternative is to treat sentences such as those in (32) as raising constructions (cf. the non-particle *be said*) separate from the tell-type control structures of (30-31).

Argument [5] concerned wh-movement. To be accounted for is the fact that (24b), with extraction of second object in a non-particle double NP construction is good, but the seemingly same extraction in (25b), with a particle construction, is ungrammatical. In the spirit of Fodor (1978), there may be an explanation of this contrast that is based in processing, rather than in principles of grammar per se, and which is consistent with a complex verb analysis of particle verbs. It is well-attested that the sentence processor fills a wh-phrase into available positions in the incoming string, on occasion erroneously anticipating the structure of the VP (see, for example, Fodor 1978; Stowe 1984). The distinction between (24b) and (25b) can be accounted for in the following way, under a complex verb and particle movement analysis. In (24b), the word following the verb hand is the indirect object John, which will alert the processor to the correct analysis, with the wh-phrase in final position; by contrast, in (25b) the verb hand plus down can be integrated into an incorrect analysis ('handed which tools down'). The difference between (24b) and (25b) can thus be accounted for as a difference in whether the correct analysis can be arrived at with only one word after the verb in hand. The erroneous analysis for (25b) may be promoted additionally by the fact that the particle down can be taken (incorrectly in the case of (25b)) as signal for closure of the verb phrase.

To summarize this section, for almost all of Kayne's arguments against a complex verb analysis, it is possible to propose an alternative analysis of the facts that is consistent with a complex verb account that conforms to the X-bar theory of word structure sketched above. In the case of the distribution of inflectional morphemes, the analysis depended on the head-modifier account of complex V structure suggested in section D. In the case of the order of constituents in dative particle constructions, our analysis depended on the existence of a particle movement rule, and makes the prediction that languages that differ in the presence vs. absence of such a rule will differ in the possibility of particle constructions with double objects; the facts of English vs. Swedish appear to support this prediction. Positing different structures to those of Kayne for infinitival complements to particle verbs allowed the control facts with respect to idioms and there to be accounted for. Wh-extraction facts were given a processing account. The right argument was not neatly dealt with, in that allowing particles to have specifier structure in verb particle constructions but not compounds requires some relaxing of the condition that constituents of words are categories at bar level 0 or lower.[15]

F. CONCLUSION

Having gone to some pains to reply to the arguments of detractors of a complex verb analysis for particle constructions in both Swedish and English, I should add that that is all that I have done. The fact that objections to a complex verb analysis can be answered does not mean that the complex verb analysis is at present strongly motivated in comparison with its competitors.

For example, an apparent virtue of the analysis above, where English and Swedish differ in the existence of a particle movement rule that raises the particle to a position where it c-commands the first object is that it allows an account of why Swedish lacks particle double object constructions. However, it appears that this lack is common to Norwegian and Danish also, languages which, like English, have particles following the object (obligatorily in the case of Danish).[16] Possibly particles that follow the object in languages that do not have double NP particle constructions are restricted in their meaning and/or by interpretive rules in such a way that they cannot legitimize the double object construction.[17] Whether this is in fact the case, or whether the absence of double object particle constructions in Scandinavian languages is merely a linguistic happenstance, seems an interesting question to pursue.

FOOTNOTES

1. Ejerhed (n.d.) uses the term compound stress to refer to the pattern of stressing on the particle. The rule that assigns stress to the particle may not in fact be part of the system of rules for stressing compound words in Swedish (below it is suggested that verb plus particle constructions are complex words but not compounds, as defined in a recent study of compounding). The stressing operation involved is more plausibly one that unifies a syntactic and/or semantic unit with end-stress, and which is responsible also for the stressing of the preposition in constructions such as (2), below (see Anward and Linell 1976, Ejerhed 1979, n.d. for discussion of a range of pertinent data; Rischel 1983 discusses similar phenomena in Danish).

2. The following spatial prepositions permit the type of semantic alternation illustrated in (2-3): bakom (behind); under (under); över (over); på (on); framför (in front of); emellan (between).

3. The speaker I have questioned in most detail in some cases finds gapping of the verb alone acceptable to some degree, in addition to gapping of both the verb and the particle. For example, her judgements were as indicated on the following sentences:

ia		Per bröt av grenen och Kalle bröt av kvisten 'Per broke off the branch and Kalle broke off the twig'
b	OK	Per bröt av grenen och Kalle ∅ ∅ kvisten
c	?/*	Per bröt av grenen och Kalle ∅ av kvisten

It is not clear what status should be assigned to examples such as (c). One possibility is that in some cases a speaker may compute (permit) a prepositional analysis as well as a particle analysis for the string, resulting in a degree of acceptability for the (c) type sentences.

4. Passivization is also a potential test of the distinction between (1) and (2-3) on the assumption that direct but not prepositional objects in Swedish may passivize (cf. Maling and Zaenen 1985, section 4). I have not checked many examples, but it appears particle constructions (defined as such by the gapping test) allow passivization and sentences of type (2), like sentences of type (3), do not freely passivize, except with the preposition <u>på</u> in passives with the verb <u>bliva</u>,

ia. Katten språng på elefanten
'The cat ran on top of the elephant'
b. Katten språng på elefanten
'The cat bumped into the elephant'
c. Elefanten blev påsprungen av katten
'The elephant was bumped into by the cat'

(Gapping of the preposition for conjoined sentences with <u>springa på</u> is rejected by everyone I have queried, with as well as without stress on the preposition).

5. Ejerhed (1979, n.d.) mentions some other points that potentially bear on the syntactic analysis of particle constructions, none of which is conclusive, as she notes. She argues against a richer analysis of the internal structure of particle constructions (in which the verb contains a VP in its internal structure, as suggested by Anward and Linell 1976) on the ground that verb particle constructions show no examples of alternation between compound stress (destressing to the left) and stress on both constitutents, parallel to contrasts found for lexicalized vs. non-lexicalized noun phrases,

Vǐta Húset 'the White House' vs.
det víta húset 'the white house'

6. Ejerhed (1979, n.d.) makes this point with examples such as få se ('catch sight of'). The sentence 'Per fick se flickan', with stress on se, has the structure [Per fick [se flickan]].

7. Another alternative is that the in-location reading is associated with a PP at a higher V-bar level than that of a PP with the towards location reading, with a correspondingly greater freedom to prepose in the in-location case. That would allow a structural basis for the assignment of stress and thematic role (section C, below).

8. A further difference not discussed here is incorporation of the particle to the left of the verb in participial forms in Swedish.

9. Numbers to the right are numbers from Kayne's article.

10. This discussion covers only the points Kayne makes against a complex V analysis (p. 125-127), and is not a complete alternative to his account of particle constructions (see footnote 15 for some discussion of points not covered in this section).

11. A different tack would be to derive examples such as (26) with right plus particle in end-position, extraposition of NP accounting for the surface order of (26). Kayne notes such examples are improved with a more complex object (Kayne's example, p. 127, 'John looked right up the information I had asked for'), consistent with the general ease of extraposition of 'heavy' NPs. This suggests a general variant on the analysis below (where particle movement raises a particle to the right of the object), in terms of intraposition and incorporation of the particle as part of the complex verb.

12. The approach here is similar to that in Lieber's 1983 analysis of compound formation in English. Possibly c-command as a condition on thematic role assignment for syntactically realized arguments will generalize constraints on compounds and particle constructions (Lieber (p. 255, fn. 6) leaves the latter out of her analysis), but I have not worked this through in detail.

13. The verbs ta på ('put on') and ta av ('take off) are exceptions to this statement,

i. Jan tog på henne skorna
'Jan put on her the shoes'

14. Kayne (p. 126) emphasizes that he is concerned with syntactic deviance of the V P NP NP sequence: "The claim is that there are no such combinations, even idiomatic, that would make [such sentences] acceptable to all, or even most, speakers". My approach differs somewhat in taking the semantically transparent cases as at the root of the ordering restriction, with the implication that the deviance of idiomatic V P NP NP sequences for many speakers is due to some kind of influence of the clear cases on judgements of the idiomatic cases.

15. The general aim of Kayne's analysis is to demonstrate that the grammar of particles will follow from the small clause analysis in interaction with general principles of grammar, and the viability of a complex V/particle movement analysis will depend not just on whether Kayne's arguments against a complex V analysis can be answered, but on how the total data covered by Kayne's analysis can be accounted for. There is not space here to detail all of Kayne's arguments. Among the arguments he gives in favor of a small clause approach to particles is that it allows a principled account of the similar behavior of V NP Adj and V NP P sequences in nominalizations (*'John's consideration of Fred honest'; *'John's calling of Bill up') and of extraction facts not dealt with in section E (particularly, the ungrammaticality of extraction of the object of a prepositional complement to NP before a particle (*'Who has the cold weather worn the sister of out?') and of the first object in double NP constructions (*?"Who should we hand (down) the tools'), both of which involve extraction from within a left-branch on Kayne's analysis (pp. 103 and 117-118). One drawback of Kayne's analysis is that verb-like aspects of the behavior of V--P sequences (the possibility of being followed by an NP and other complement types (p. 107-8 and 128)) and the possibility of nominalization with the particle in immediate post-verbal position ('John's calling up of Bill') require introduction of a special mechanism of theta role-percolation to distinguish particle and adjectival small clauses (p. 128-130). Gapping is not discussed in Kayne's article, and it is a point in favor of a non-small clause analysis that Kayne's analysis, in which the particle and NP comprise an immediate constituent of a projection of V, incorrectly predicts gapping of a verb plus particle (example 9b in the text) to be ungrammatical.

16. I have not checked this for predicative constructions in Danish. Herslund (n.d.) notes the absence of double NP particle datives in Danish, with the exception (fn. 8) of <u>give tilbage</u> ('give back'),

i. De gav ham hans penge tilbage
'They gave him his money back'

observing that this use of <u>tilbage</u> does not freely extend to other double-object taking verbs. From my informants' judgements <u>ge (NP) (NP) tillbaka</u> is a similarly restricted exceptional construction in Swedish. (See also Kayne 1985, p. 120 on the fact that <u>back</u> in English is acceptable in final position after two objects, unlike particles such as <u>down</u>).

17. The same structural position may be more restricted in interpretation in one language than another. Åfarli (1985) argues end-placed particles in Norwegian participate in a causative rule system. The following examples illustrate that end-placed particles in English also occupy a position associated with result or cause. In (ia) sweaty may have a result interpretation; in (ib,c) sweaty may only have a predicative interpretation (with respect to either John or the pigs),

ia. John drove the pigs sweaty
b. ?John drove in the pigs sweaty
c. John drove the pigs in sweaty

It is not plain to me whether the preference for middle position of the particle in adjectival constructions (noted by Jackendoff 1977, p. 67) has the same source as the preference for middle position in double NP constructions. One possibility concerning the deviance of semantically transparent examples such as (ib) is that there is a tendency to misconstrue the final adjective as a result, which will be inadmissible since result is preempted by the particle.

ACKNOWLEDGEMENTS

This work was supported by funds from the Swedish Institute and the Graduate School of the University of Wisconsin-Madison (project no 170668 to Helen Goodluck). Elisabet Engdahl and Christer Platzack among others provided helpful references and comments and Anne-Christine Bredvad-Jensen was a particularly patient informant.

REFERENCES

Anward, J. and P. Linell. 1976. Om lexikaliserade fraser i svenskan, Nysvenska Studier, Årg 55-56, 77-119.

Bolinger, D. 1971. The Phrasal Verb in English, Cambridge, MA: Harvard University Press.

Ejerhed, E. 1979. Verb-partikel-konstruktionen i svenska: syntaktiska och semantiska problem. O. Josephson, H. Strand, N. Westman (eds). Förhandlingar vid sammankomst för att dryfta frågor rörande svenskans beskrivning 11, Stockholm.

Ejerhed, E. n.d. Swedish verb-particle constructions: semantic and syntactic problems. Ms.

Emonds, J. 1972. Evidence that indirect object movement is a structure preserving rule. Foundations of Language, 8, 546-61.

Fodor, J. 1978. Parsing strategies and constraints on transformations. Linguistic Inquiry, 9, 427-453.

Gruber, J. 1976. Studies in Lexical Relations. In Lexical Structures in Syntax and Semantics. North Holland Publishing Company, Amsterdam.

Herslund, M. n.d. The double object construction in Danish. Ms. University of Copenhagen.

Jackendoff, R. 1977. $\bar{X}$ Syntax: A Study of Phrase Structure, Cambridge, MA: The MIT Press.

Kayne, R. 1984. Connectedness and Binary Branching. Foris Publications, Dordrecht.

Kayne, R. 1985. Principles of particle constructions. In Grammatical Representations, edited by J. Guéron, H.-G. Obenauer and J.-Y. Pollack. Foris Publications, Dordrecht.

Lieber, R. 1983. Argument linking and compounds in English. Linguistic Inquiry, 14, 151-85.

Maling, J. and A. Zaenen. Preposition stranding and passive. Nordic Journal of Linguistics, 8, 197-209.

Neijt, A. 1979. Gapping: A Contribution to Sentence Grammar. Foris Publications, Dordrecht.

Rischel, J. 1983. On unit accentuation in Danish - and the distinction between deep and surface phonology. Folia Linguistica, XVII/1-2, Special Issue "Prosody".

Selkirk, L. 1982. The Syntax of Words. Cambridge, MA: MIT Press.

Stillings, J. 1975. The formulation of gapping English as evidence for variable types in syntactic transformations. Linguistic Analysis, 1, 247-74.

Stowe, L. 1984. Models of Gap Location in Human Parser. University of Wisconsin dissertation, distributed by the Indiana University Linguistics Club.

Åfarli, T. 1985. Norwegian verb particle constructions as causative constructions. Nordic Journal of Linguistics, 8, 75-98.

A visit to China 29.4. - 7.6. 1986

Eva Gårding

In the fall of 1985 I was invited to lecture and do research at the Institute of Acoustics of Academia Sinica, Beijing. IVA, the Swedish Academy of Engineering Science, accepted me in their exchange program with China. Lund University gave me a travel grant from Elisabeth Rausing's foundation. What follows is a translation of my report to IVA and Lund University.

The Institute of Acoustics, which was founded in 1964, has about 200 employed researchers and comprises 13 laboratories with activities ranging from signal processing, underwater acoustics, room acoustics, geo acoustics etc. to the acoustics of speech. The activities include basic research with applications as well as supervision of master's dissertations and doctoral dissertations. The head of the speech acoustics laboratory is Professor Jialu Zhang.

My invitation to China was motivated by a collaboration between the Institute of Acoustics and the Phonetics department in Lund which started during the spring of 1983 when Professor Zhang was a guest researcher at a project dealing with some non-European languages, among those Chinese, and sponsored by the Research Council. The interest in intonation among speech engineers has to do with the present development of text-to-speech and speech- to-text systems. Zhang had become interested in the model for analysis and synthesis of Swedish intonation which had seemed applicable also to Chinese.

In the traditional view of intonation, words with their tones or accents are regarded as basic primary units and the sentence intonation expressing statement, question, etc. as secondary. The difficulty will then be to explain the varying manifestations of accents and tones that occur in practice. In our model this order is reversed. Accents and tones are superimposed on a larger movement determined by sentence intonation. In this way the different manifestations of accents and tones in the sentence can be derived by rather simple rules. The model is more economical than the traditional one and has a natural relation to speech psychology.

At the Acoustics Institute I gave a series of lectures, entitled "Intonation across several languanges". They were organized according to the communicative functions of intonation, i.e. to mark lexical units, to partition speech into phrases, to show the focus and syntactic type of the sentence and the emotional attitude of the speaker. The examples, taken from Swedish, French, English and Chinese, were intended to show that if lexical markings are factored out, the intonation expressing the other functions of communication is remarkably similar.-- Some days were spent on sketching an outline for future collaboration on intonation.

I also gave two lectures at the spring institute of modern phonetics in Tianjin for university teachers from all over China. The titles were "The structure and manifestations of intonation" and "Learning and teaching problems in the light of a model for prosody". The first of these lectures was later given at Beijing University. There I could also test, with the aid of professor Zhang, stimuli which hade been prepared in our laboratory by Shi Bo, a research student from the Acoustics Institute who has a grant from the Swedish Institute. The work is part of her research project whose aim is to quantify our qualitative model.

I lived in comfortable quarters at the Friendship Hotel where Academia Sinica houses her guests and provides them with a car and driver. Sightseeing and evening entertainment were part of the program. In the restaurant of the hotel I met scientists from many countries invited by Academia Sinica. The concensus was that China's efforts to improve research and higher education were promising. These efforts were a standing theme in the English language newspaper "China Daily" and it was emphasized there that exchange programs for research are essential for the economic development of the country.

Together with my husband Lars Gårding, who had been a guest lecturer at Nankai University in Tianjin, I was invited by its mathematical institute to a two week tour of Southwest China with visits to the universities of Kunming, Chengdu, Qungqing and Wuhan.

I was impressed by the kindness and care shown to me and the interest in my scientific work. Professor Zhang emphasized more than once the importance of contact between speech technology and speech science.

I should like to thank in particular my host, professor Zhang, and his collaborators at the speech acoustics laboratory, Mr Qi and Mr Lu as well as the present and former director of the Acoustics Institute, Professor Guan and Professor Maa for all favours rendered to me. I am also grateful to Professor Xing and Mr Shi of the Department of Chinese Linguistics at Nankai University for their kind reception and hospitality.

Production and perception of phrases in some Nordic dialects

Eva Gårding and David House

Paper presented at the Sixth International Conference of Nordic and General Linguistics. August 18-22, 1986, Helsinki.

Introduction

Intonation helps us group words into phrases and phrases into sentences. This grouping is achieved with the aid of both connective and demarcative features. I have listed some of them here.

Connective signals may be:

(1) similarity of the elements of the group

(2) recurrent special patterns, rhythmic or melodic /1/

(3) tonal links between the elements (junctures)

Demarcative signals may be:

(1) breaks of similarity, for instance by the introduction of new elements like pauses, internal junctures, etc.

(2) special markers used to denote the beginning or end of a group.

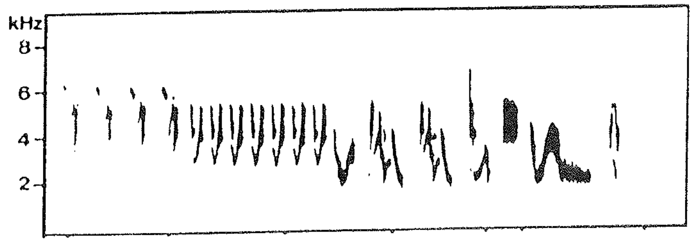
(3) boundaries between recurrent patterns

Phrasing occurs not only in speech but in all sorts of time-ordered sequences, in bird song and music to give you two examples. Even here it is meaningful to speak of connective and demarcative signals /2/.

Take the song of the chaffinch for instance. It repeats the same pattern over and over again. As you can see in the spectrogram (Fig.1), the pattern has different groups connected by similarity of the elements, an introductory ti,ti,ti followed by a trill which changes to a slower trill which ends with a special marker, the concluding tiuiu. Note that the entire pattern is falling. The part of the song that you can see in the spectrogram is called a strophe by ornithologists (e.g. Bergmann & Helb 1982, from which the spectrogram is borrowed) and the groupings of similar elements are called phrases.

The next example is from Beethoven's bagatelle opus 33 No.6 (Fig. 1). In this line, up to the point which has been marked by an arrow, there are two groups of equal duration. The first group is falling with two markers, an introductory onset and a final strong long note. Beethoven marks this group with a slur. The second group starts with an onset and ends with a final marker consisting of three notes. This group is marked by three slurs, indicating three subgroups, one over the rising part and two over the final falling parts. The similarity with the intonation and rhythm of spoken sentences may be the reason why Beethoven added Con una certa espressione parlante to this piece. The grouping of music by melodic and rhythmic means is commonly referred to as phrasing. It is a very essential part of music and all important in its execution.

After this general introduction, I will only talk about phrasing in speech. Our talk is a report on work in



FRINGILLA CEOLEBS

BEETHOVEN OP. 33 NO. 6

Fig. 1 Examples of phrasing in bird song and music progress. It aims at studying the production and perception of phrases in different prosodic systems. For today's talk we have chosen Swedish, Danish and Finnish /3/. I will talk about the production of phrases in a simple grouping experiment and make a summary comparison with real speech and David House will talk about some perception experiments.

Let me start with a rough definition of a prosodic phrase. A prosodic phrase is a part of an utterance in which accents or tones are coordinated in a unifying intonation pattern.

A grouping experiment

I will first describe a simple experiment in which speakers of different Swedish dialects and of Danish and Finnish were asked to say a series of fives, we called it a telephone number, as an ungrouped sequence and in groups of two and three and three and two, all in declarative intonation. The speakers used their respective languages and the Finns the colloquial monosyllabic form viis. With a sequence of numbers we meant to have stable experimental conditions with controlled influence from semantics and syntax.

You will now hear recordings of some of our informants and at the same time we will show you the corresponding pitch curves. The figures are practically self-explanatory. Just a few comments are needed. Look at Figure 2, for instance. The thick lines mark the intonation over the vocalic part of the syllable, the thin lines mark the voiced consonants and the dotted lines the voiceless ones. These dotted lines indicate the intonation movement masked by the voiceless segments. The parallel broken lines which enclose the curves are

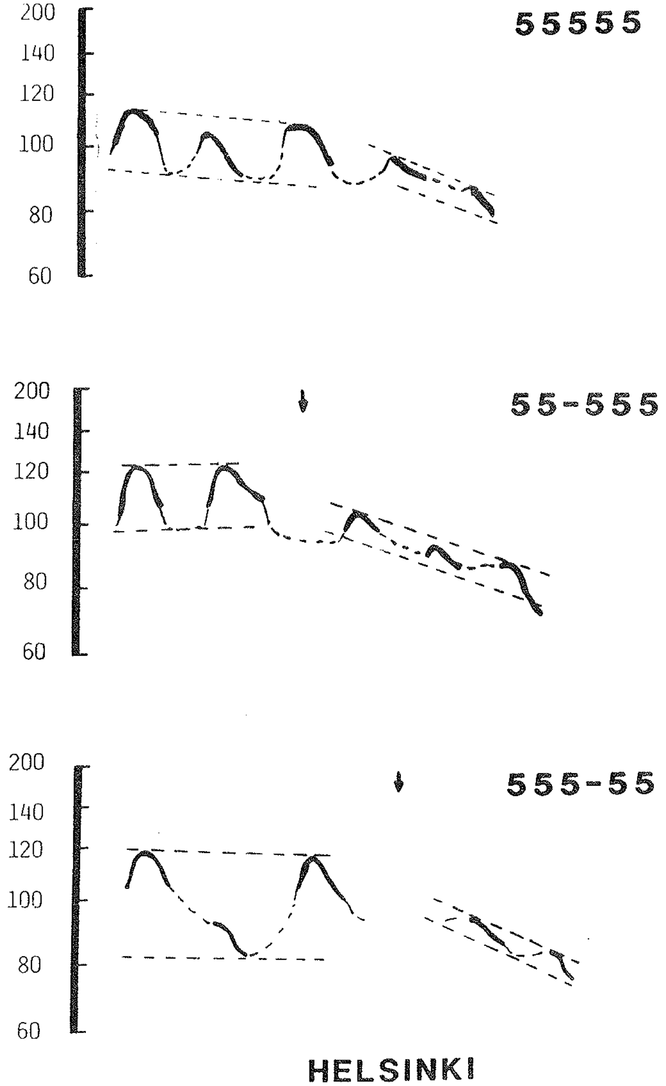


Fig. 2 Pitch curves from groupings of fives

auxiliary lines. We call them grids. They make it easy to see the general direction of the intonation and the point in time where there is a change of direction. This point is called a pivot and it is marked by an arrow. The width of the grid is correlated to stress. The wider the grid, the stronger the stress. A grid is convenient because it gives a combined measure of stress and pitch /4/.

(Here followed a demonstration of speakers from Helsinki, Stockholm, Gothenburg, Malmö and Copenhagen, accompanied by Figures 2,3,4,5 and 6.)

It is obvious from what you have just heard and seen that the speakers use both general and language-specific prosodic features to perform the required groupings. Let us look more closely at Figure 7 where I have combined the groupings performed by speakers from Helsinki and Stockholm.

The Finnish speaker has a wide, rather level grid for the first phrase and a narrow falling grid for the second. In the wide grid we see a special pattern in the group of three, which we can label High-Low-High. This gives it a trough-like tonal shape. The second group has a similarity pattern with three individual falls.

The Stockholm speaker has rising-falling grids to mark the two phrases. Like the Finn, she has a trough-like pattern in the first group but unlike him, she has a special marker, the rising movement to the high phrase accent for the non-terminal phrase and the same high phrase accent combined with a terminal fall to mark the end of the final group. The high phrase accent is typical of Central Swedish dialects (Bruce 1977). All in all, this is the tonal manifestation of an iambic or anapestic rhythmic group compared to the Finn's evenly spaced viises forming spondaic groups.

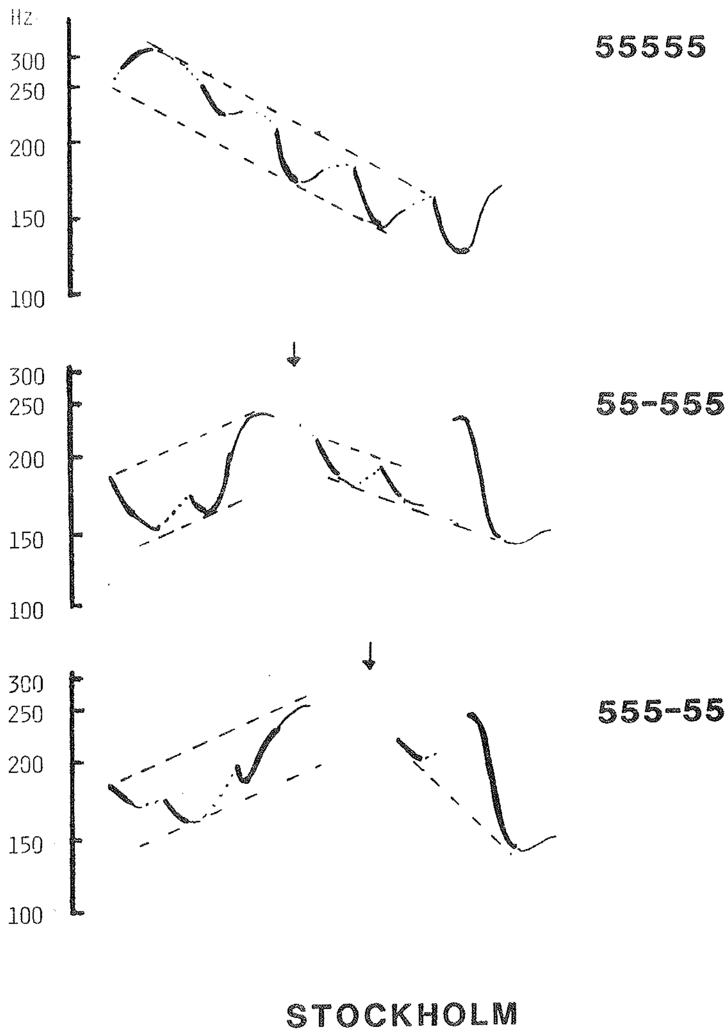


Fig. 3 Pitch curves from groupings of fives

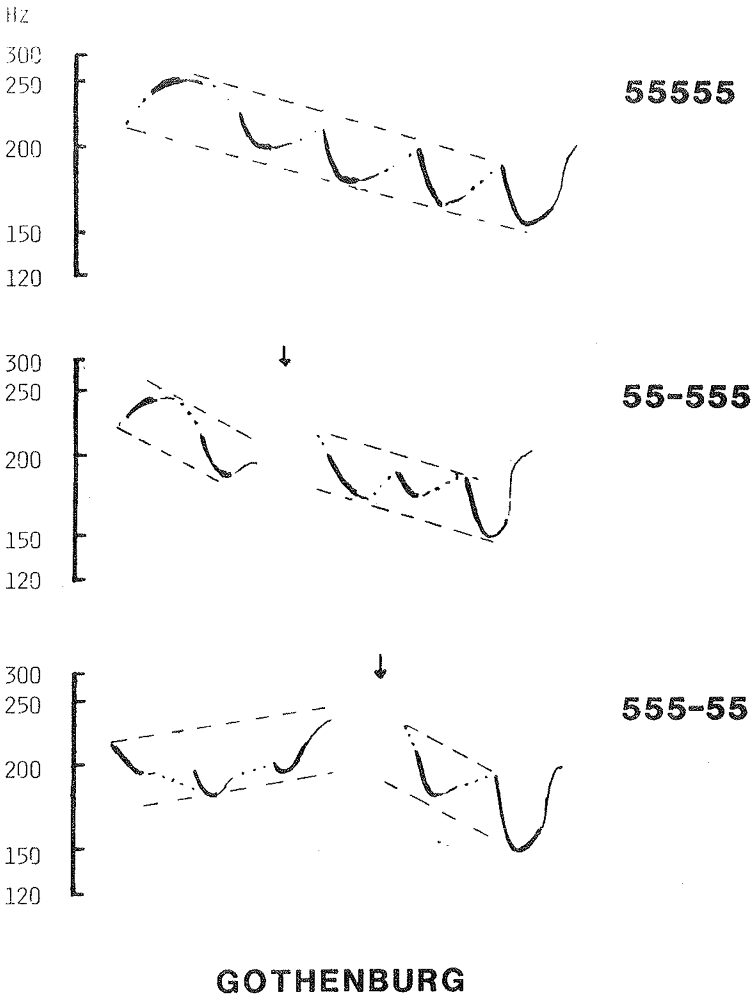


Fig 4 Pitch curves from groupings of fives

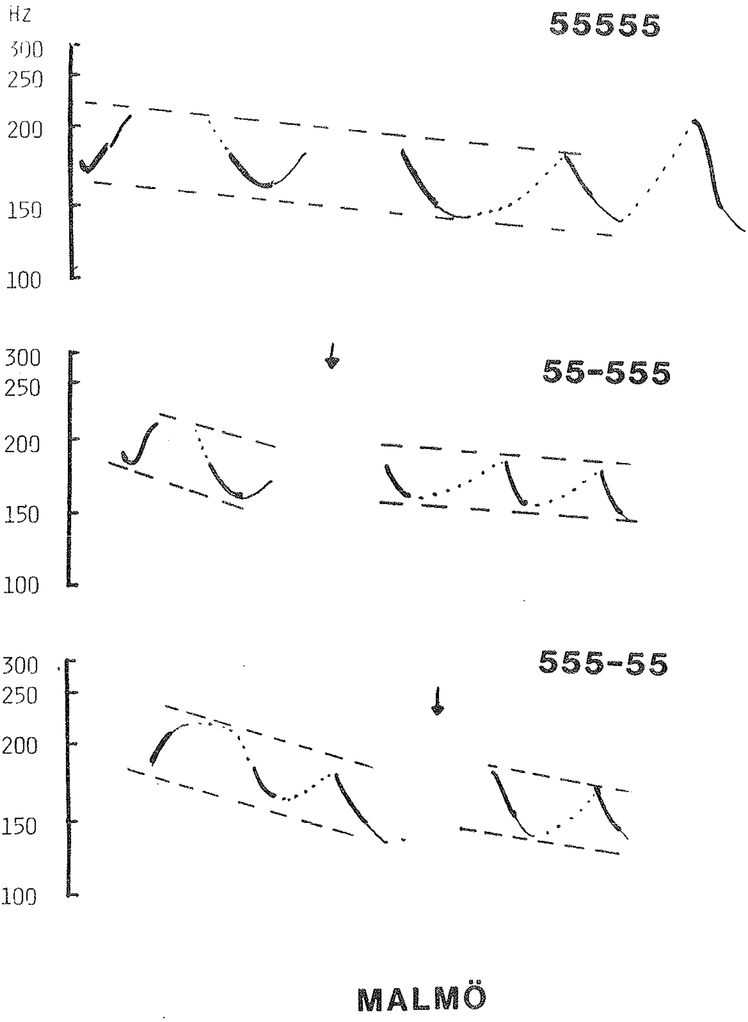
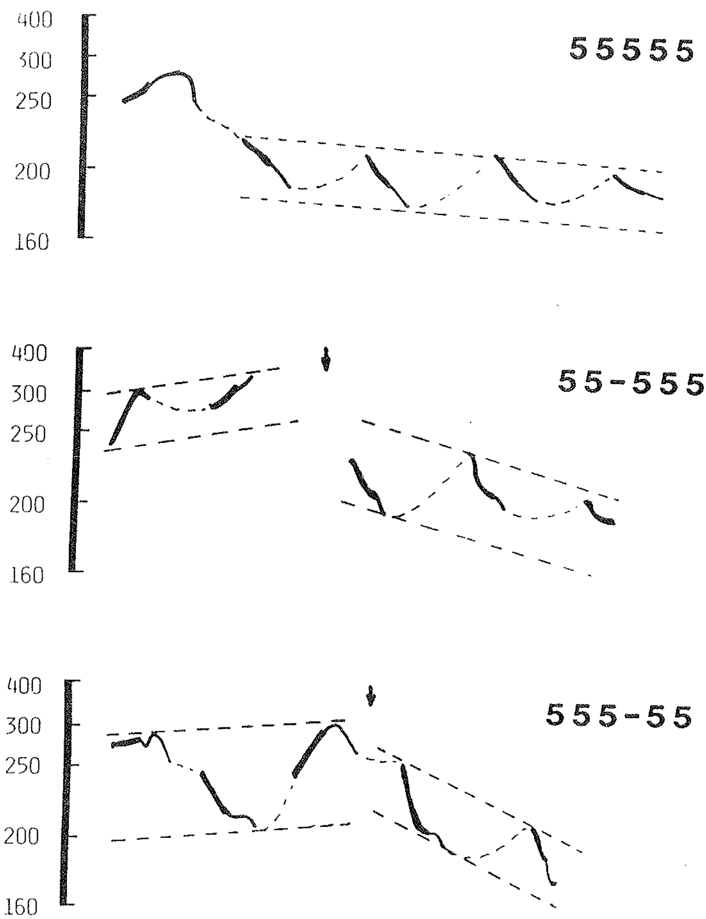


Fig. 5 Pitch curves from groupings of fives



COPENHAGEN

Fig. 6 Pitch curves from groupings of fives

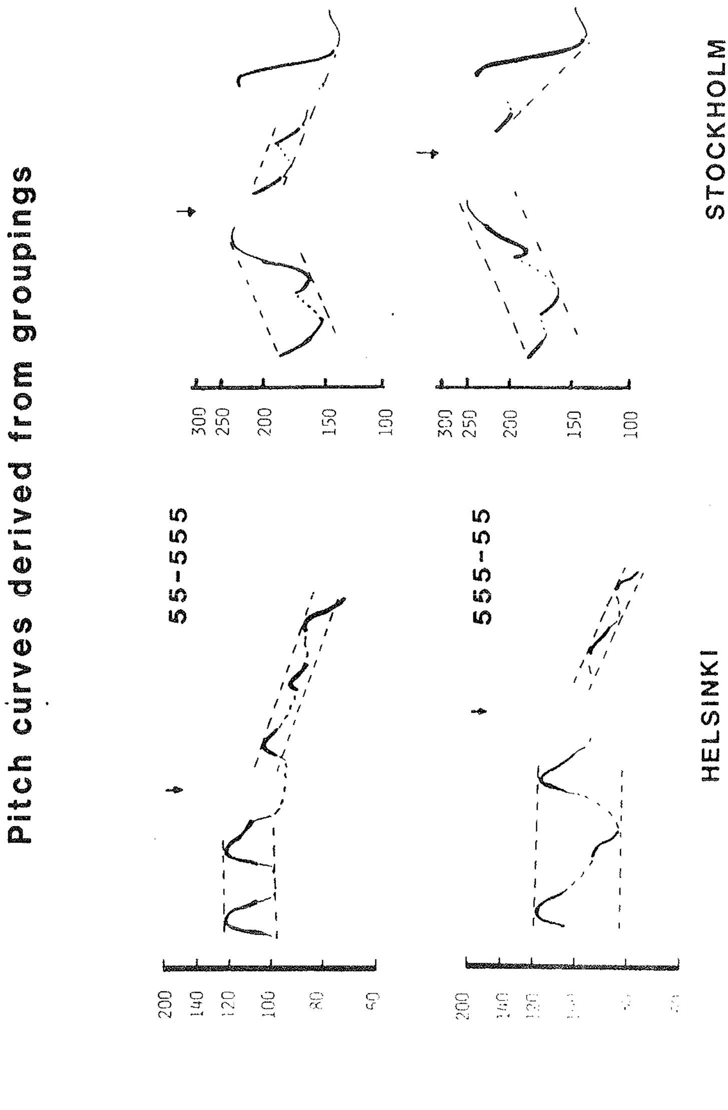


Fig. 7 Comparison of prosodic features used in groupings

These two speakers gave an example of a language-specific difference in grouping habits.

The Stockholm and Malmö speakers are examples of dialect-dependent variation (Fig. 8). The high phrase-final accent of Stockholm is in Skåne manifested as a low accent which is one of the reasons why these two dialects sound so different. Another reason is the Skåne speaker's spondaic rhythm which, not surprisingly, makes it more like Danish.

There are also examples of free variation in our material with some speakers using different strategies in their repeated renderings of the groups. Figure 4 gives an example of a Gothenburg speaker who wavers between two consecutive falls and a rise-fall for the groupings of the number.

Now let us look for some general principles present in all of the investigated dialects.

As demarcative signals all the speakers use pivots, manifested as pauses or a change of the grid range or grid direction. As connective signals all the speakers, irrespective of language and dialect, use falling similarity patterns in the second group, carrying declarative intonation, and special patterns in the first group (Fig. 9). The shape of the special pattern is determined by the terminal part of the group, a rise tends to be preceded by a fall and vice versa. The shape of the terminal part is dialect dependent. In groups of three, the mid accent is weakened and lowered which is a stable contribution to the pattern.

On the whole, this material reveals a very strict preplanning. When the last group is falling, the first group is level or rising. When the last accent of a group is

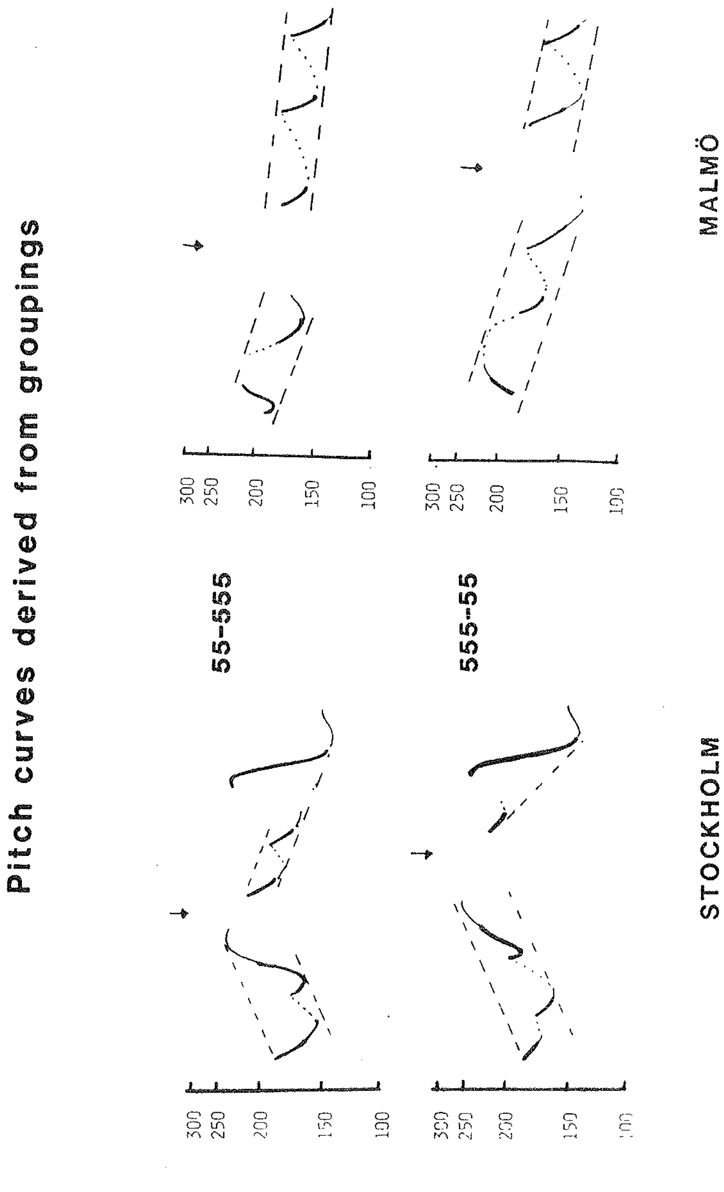


Fig. 8 Comparison of prosodic features used in groupings

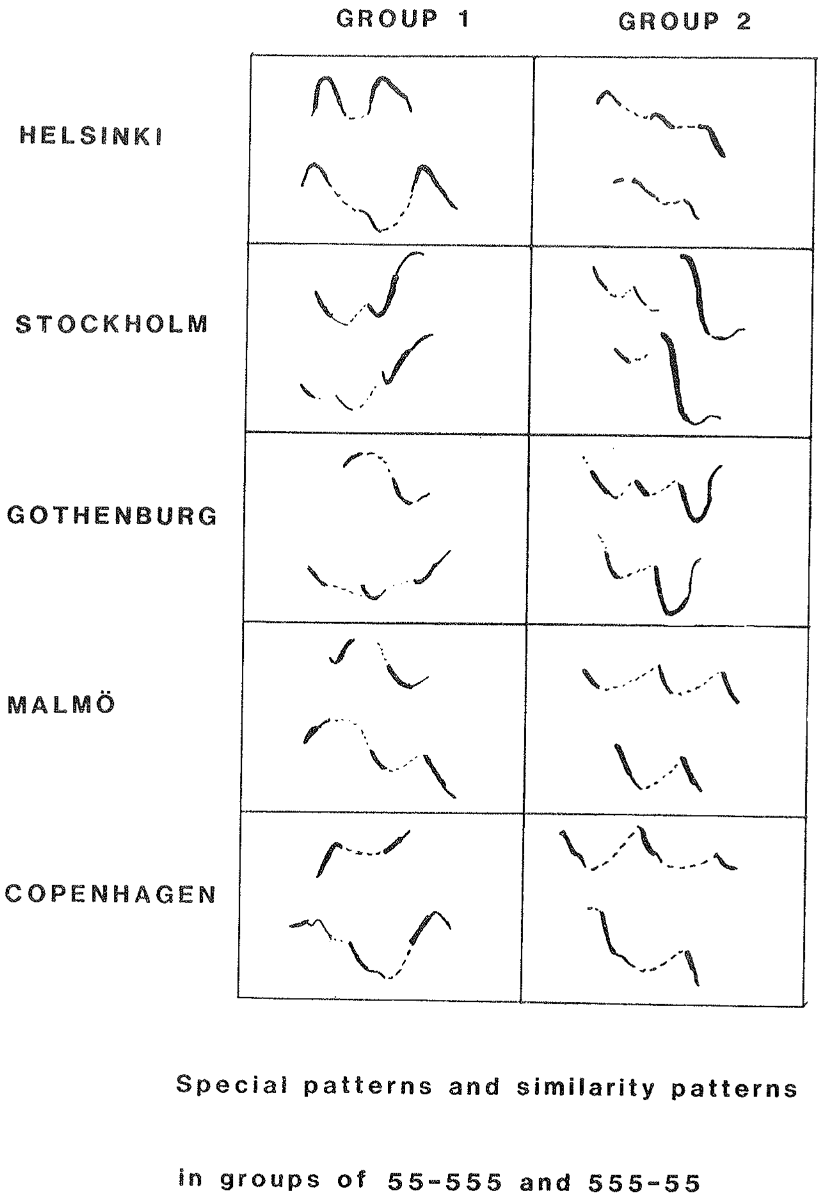


Fig. 9 Special patterns and similarity patterns in groups of 55-555 and 555-55

rising, the preceding ones are falling, etc.

Phrasing in real speech

For this short talk we have decided to give preliminary answers to two questions. One has to do with connective signals, more precisely, the order-bound use of similarity patterns and special patterns. Is there any correspondence to this in real natural speech?

To find an answer to this question, we used a recording of a spontaneous conversation between Stockholm people, moderated by Bengt Loman in Malmö in 1966. (Gårding 1967a is an analysis of this material).

The answer is YES. Similar things do occur. Factors that determine the choice between special patterns and similarity patterns seem to be semantic importance and semantic coherence. In fluent narrative style the great majority of sentences seem to contain a focussed part, the rheme, which often comes early in the sentence. It is phonetically marked by a wide grid and its elements are closely connected in a special pattern. The rest of the sentence is backgrounded with similarity patterns in more or less compressed grids.

The other question that we asked of the material has to do with demarcation. What about phrase boundaries in real speech? Here one has to distinguish between precise boundaries and boundary regions. Precise boundaries are possible to detect in the acoustic record only when two accented syllables meet with an internal juncture between them /5/. In other cases, precise boundaries are replaced by boundary zones consisting of consonants or unaccented syllables of which some may be enclitic and others proclitic. Their intonation forms a bridge between the adjoining accented syllables and this bridge is constructed according to the principle of the shortest way (Gårding and House 1985).

The implication is that in this case the precise boundaries cannot be detected by acoustic criteria, quite simply because they are not there. What listeners do, then, if they are asked to segment such a sentence into phonetic phrases, is to let syntactic and semantic criteria guide their segmentation (Gårding 1967a p. 51 ff.).

With this comment we have approached the perceptual part of our talk and David House will take over.

Perception of grouping

I am going to talk about some perception experiments which tested the relevance of the features we have observed. In particular we are interested in the relevance of the grid and the pivot as connective and demarcative signals /6/. For the tests 34 different synthetic stimuli consisting of sequences of fives were randomized and presented to 20 Swedish listeners according to standard procedure. Each stimulus received 100 responses. Listeners were asked to judge if the stimuli were grouped 2+3 or 3+2. As test stimuli we used manipulated variations of a natural Scanian <u>fem</u> 'five'. Note that pause is not one of the variables and that only tonal patterns are represented in the material.

Figure 10 shows a few examples of the stylized tonal contours used as stimuli. The results are given as percent "correct" response where "correct" means "expected" or "anticipated". For example, we expected stimulus 1, designed

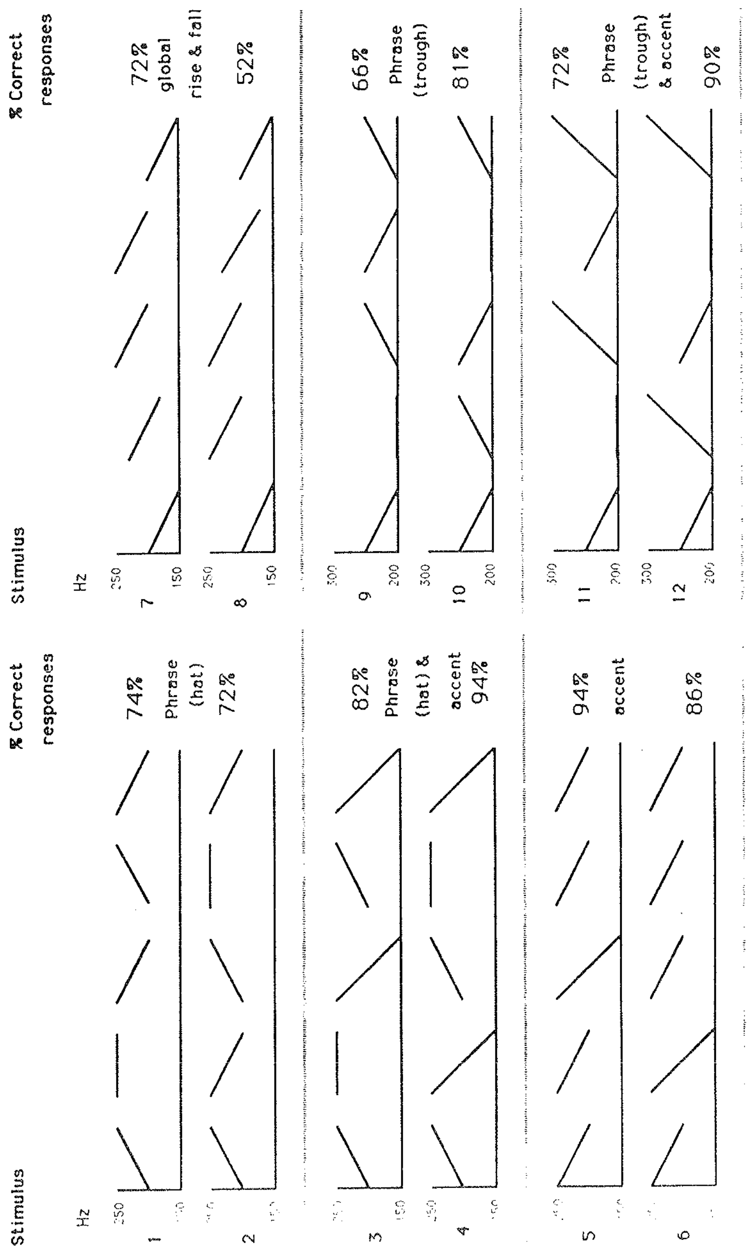


Fig. 10 Examples of stylized tonal contours used as stimuli

to test a hat-like pattern as a connective signal to be grouped as 3+2 and, indeed, 74% of the responses were 3+2. Each stimulus has a counterpart, in this case stimulus 2 with the expected grouping 2+3. Here 72% of the responses were 2+3.

The next set of stimuli were designed to test the combination of a hat-like pattern and a large tonal movement as a demarcative signal. When both the connective and demarcative signals were used, the results improved considerably. Stimulus 3 now received 82% and stimulus 4 92% of the votes.

The following set shows the power of the special marker alone. Here we get 94% and 86%. More examples are rising and falling patterns and a trough-like pattern as connective signals. Finally, the trough-like pattern is combined with a special marker which again improves responses.

(Here the 12 sample stimuli were played.)

Our test shows that similarity of consecutive elements is a good connective signal. So are recurrent patterns (1,2 hats, 9,10 troughs, and 7 a rise and fall).

The special markers, the large-size tonal movements seem to be the best demarcative signals. It is interesting that these markers, both the falling and rising ones, are judged to be final elements. One possible explanation is that in Swedish a large final tonal movement is a characteristic feature of phrase stress or juncture. A break of direction or resetting of a grid is also a good demarcative signal. One exception is number 8 where this demarcative signal seems to be in conflict with a connective signal consisting of three very similar elements across the intended border.

The prosodic phrase in our model of intonation

Let me now come back to the definition of prosodic phrase. It is natural to apply the term prosodic phrase to each group of our two-group sentences. Our original definition of prosodic phrase in this talk is: a part of an utterance where accents and tones are coordinated in a unifying intonation pattern. We can now make the definition more precise: a prosodic phrase is a part of an utterance which is connected by a special pattern and bounded by pivots. A pivot can be a special marker or a change of direction or range of the grid. This is the definition used in our model of intonation/7/. It provides a useful tool for analyzing the aspects of intonation which help us group words into phrases and phrases into sentences, both in a Nordic and a general perspective.

SUMMARY

The grouping of time-ordered sequences uses both connective and demarcative features. Connective signals are similarity of the elements of the group and recurrent special patterns. Demarcative signals are e.g. breaks of similarity and special markers at the beginning or end of a group. The relevance of such signals for speech was first tested in a grouping experiment with numbers, 55555 contrasted with 55.555 and 555.55. The experiment was carried out with speakers representing Finnish (Helsinki), Swedish (Stockholm, Gothenburg and Malmö) and Danish (Copenhagen).

A common feature for all these dialects is that groups (phrases) are coordinated in a common unifying intonational

frame (tonal grid) and that discontinuities in this frame (pivots) correspond to phrase boundaries.

Another common feature in our material is the order-bound use of similarity patterns consisting of individual falls in the final group carrying declarative intonation, and the preference for a special pattern in the initial group. In initial groups of three, the mid accent is generally reduced.

The shape of the special pattern is dependent on the phrase-final accent which varies with dialect.

The number material suggests that there is a high degree of preplanning at two levels. At the sentence level the intonation of the last phrase seems to determine that of the first, and at the phrase level the shape of the last accent of the group determines the movements of the preceding ones. The governing principle seems to be one of contrast. A communication-carrying fall, as in a declarative sentence, is sharpened by a preceding rise and vice versa.

There is a certain amount of free variation for some of the speakers.

Features comparable to those found in the number material are observed in a recording of spontaneous Swedish speech. In a narrative style most sentences contain a focussed part marked by a wide grid and with its elements connected in a special pattern. The rest of the sentence is backgrounded with similarity patterns in compressed grids. As a rule, syntactic boundaries in a series of unaccented syllables are phonetically unmarked.

A complementary investigation of the perceptual relevance of connective and demarcative signals demonstrated that listeners can use the corresponding acoustic correlates, grids and pivots, as cues to divide a series of numbers into two groups. Duration was not a variable in these experiments as the duration of each synthesized stimulus and the interval between them were always the same. The most powerful cue was a large tonal rise or fall which was judged to be a final element in a group.

NOTES

/1/. The special patterns are determined by phonological rules. In function and partly in form they are comparable to the results of Chinese sandhi rules.

/2/. In his classification of phonetic phenomena, Elert notes the importance of distinguishing connective features from demarcative ones (Elert 1970 p. 20).

/3/. Earlier reports are Gårding and House 1985 and House and Gårding 1986.

/4/. The notions of tonal grid and pivot with examples from several prosodic systems are explained in Gårding 1984 and Gårding 1985.

/5/. In Gårding's doctoral dissertation, internal juncture is defined as a marked syllable boundary in a phrase (Gårding 1967b p.33). Acoustic correlates are glottal stops or lengthening of the segments at the boundary.
The first comprehensive analysis of internal juncture had been made by Lehiste for English (1960).

/6/. That intonation by itself is a cue to grouping is already known from results reported in the literature (Collier & t'Hart 1975, Lehiste 1973).

/7/. There are at present different versions of this model. See e.g. Bannert 1984, Bruce 1977, 1982, Bruce & Gårding 1978, Gårding 1979, 1981, 1983.
An early presentation of the generative part of the model was given at the Second Conference of Nordic Languages and Modern Linguistics (Gårding 1973).

REFERENCES

Bannert R. 1984. Towards a model for German prosody. Working Papers 27, Department of Linguistcs, Lund University: 1-36

Bergmann H.-H.& H.-W. Helb. 1982. Stimmen der Vögel Europas. BLV Verlagsgesellschaft mb H. München

Bruce G. 1977. Swedish word accents in sentence perspective. Travaux de l'Institut de Linguistique de Lund XII. Gleerups, Lund

Bruce G. 1982. Experiments with the Swedish intonation model. In Preprints of Working group of intonation, Int. Congress of Linguistics, eds. H. Fujisaki & E. Gårding, Tokyo: 35-45 and Working Papers 22, Department of Linguistics, Lund University: 51-116

Bruce G. 1984. Aspects of Fo declination in Swedish. Working Papers 27, Department of Linguistics, Lund University: 51-64

Bruce G.& E. Gårding. A prosodic typology for Swedish dialects. In Nordic Prosody I, eds. E. Gårding, G. Bruce & R. Bannert. Travaux de l'Institut de Lund XIII: 219-228

Elert C.-C. 1970. Ljud och ord i svenskan. Almqvist & Wiksell. Stockholm

Gårding E. 1967a. Prosodiska drag i spontant och uppläst tal. In Svenskt talspråk, ed. G. Holm. Almqvist & Wiksell. Stockholm: 40-85

Gårding E. 1967b. Internal Juncture in Swedish. Travaux de l'Institut de Phonétique de Lund VI. Gleerup, Lund.

Gårding E. 1975. Toward a prosodic typology for Swedish dialects. In Proc. Nordic and Modern Linguistics II, ed. K.-H. Dahlstedt. Almqvist & Wiksell International, Stockholm: 466-475

Gårding E. 1978. Sentence intonation in Swedish. Phonetica 36: 207-215

Gårding E. 1981. Contrastive Prosody: A model and its application. Studia Linguistica 35: 146-166

Gårding E. 1983. A generative model of intonation. In Prosody: Models and measurements, eds. A. Cutler & R. Ladd, Springer: 11-25

Gårding E. 1984. Comparing intonation. Working Papers 27, Department of Linguistics, Lund University: 75-100

Gårding E. 1985. In defence of a phrase-based model of intonation. Working Papers 28, Department of Linguistics, Lund University: 1-18

Gårding E., A. Botinis & P. Touati. 1982. A comparative study of Swedish, Greek, and French intonation. Working Papers 22, Department of Linguistics, Lund University: 137-152

Gårding E. & D. House. 1985. Frasintonation, särskilt i svenska. In Svenskans beskrivning 15, eds. S. Allen et al. Göteborgs universitet, Göteborg: 205-221

House D. 1985. Sentence prosody and syntax in speech perception. Working Papers 28, Department of Linguistics, Lund University: 91-107

House D. & E. Gårding. 1986. Phrasing in some Scandinavian Dialects. Paper presented at the fourth Nordic Prosody

Conference in Middelfart, Denmark.
In preparation.

Lehiste I. 1960. An acoustic-phonetic study of internal open juncture. Phonetica 5. Suppl.

Lehiste I. 1973. Phonetic disambiguation of syntactic ambiguity. Glossa 7: 107-122

Loman B. 1966. A recording of a spontaneous conversation between Stockholm speakers. 2.5 hours on tape, in care of E. Gårding.

Final report: Phonetic analyses of some non-European languages (LUCLA)

Eva Gårding, Mona Lindau, Kjell Norlin and Jan-Olof Svantesson

The acronym stands for a phonetic collaboration between Lund University and UCLA on a project sponsored by the Swedish Council for Research in the Humanities and the Social Sciences. The aim of the project was to provide phonetic data and analyses of Egyptian Arabic, Hausa, and Standard Chinese. Mona Lindau was responsible for Hausa, Kjell Norlin for Arabic, and Jan-Olof Svantesson and Eva Gårding (not paid) were responsible for the Chinese part. The Research Council granted additional financial support for Jialu Zhang (Acoustics Institute, Academia Sinica, Beijing) who spent three months at the phonetics department in Lund. We have also enjoyed the collaboration of Paul Kratochvil (Faculty of Oriental Languages, Cambridge) and Kristina Lindell and Magnus Nordenhake (Department of East Asian Languages, Lund University). Peter Ladefoged (Phonetics Laboratory, UCLA) provided much technical support, computer programs, and encouragement. Taghrid Anbar (Cairo University) spent three months at Lund with a scholarship from the Swedish Institute, working on the pedagogical application of Arabic.

The goals of the project in the grant proposal were stated as follows:

1. Analysis of speech sounds and prosody.
2. Cross-linguistic comparisons.
3. Pedagogical applications

A major principle of the project was to collect data of the selected languages in a uniform manner, so that valid cross-linguistic comparisons can be made. The data collection consists of tape recordings of several speakers for each language so that the results are representative of the language, not just of a single speaker. Ten speakers of Kano Hausa were recorded in Nigeria, eight speakers of Cairo Arabic were recorded in Lund, and six speakers of Standard Chinese were recorded in Lund, Stockholm, and Beijing. These data were then subjected to acoustic analyses which have been presented as papers in journals (see Bibliography). The major results of the cross-linguistic studies are summarized below.

Stops

(The consonant systems of the three languages are shown in the appendix.)

All three languages have labial, dental/alveolar, and velar stops that are common in the languages of the world. Chinese has two series of stops, voiceless unaspirated and voicelesss aspirated. Hausa and Arabic have voiced and voiceless stops and in addition they have a third, more unusual, stop series. In Hausa this third series is described as "glottalized", in Arabic the third series is pharyngealized.

The durations of the closure and the aspiration in the stops were measured for all three languages. Data for closure duration and aspiration for Swedish from Löfqvist (1976) was used for comparison. Figure 1 shows a diagram of the mean values of these durations. The duration of the aspiration in Hausa, Arabic and Swedish are 30-60 milliseconds. These values are typical of languages with a voiced/voiceless contrast. In Chinese both series are voiceless and the burden of contrast lies in the aspiration. The duration of the aspiration in the so-called unaspirated stops is 5-30 ms. The aspirated stops have a very long aspiration (about 100 ms). The closure durations are significantly different in the three languages, being the longest in Swedish, and the shortest in Hausa. The Hausa stops are thus generally shorter than Swedish stops. There was not a good correlation between closure duration and aspiration, so the relationship between closure duration and aspiration appears to be language specific.

The "glottalized" stops in Hausa are usually written /ɓ/, /ɗ/ and /k'/. The main characteristic of these stops in Hausa is a laryngealized phonation lasting throughout the stop and well into the following vowel for the voiced /ɓ/ and /ɗ/, and following the release of the voiceless /k'/. In addition, the voiced /ɓ/ and /ɗ/ are either implosive or weakened to [w] and [ð], respectively, and /k'/ is ejective. Just as the regular stops, these laryngealized stops are relatively short compared to implosives and ejectives in other languages. Mainly due to their laryngealized phonation, the implosive Hausa stops differ considerably from implosives and ejectives in neighbouring, unrelated Niger-Congo languages. This supports the notion that the glottalized consonants in Hausa are indigenous rather than borrowed (Lindau 1984).

Arabic plain and pharyngealized voiced stops show no significant difference in duration or waveform, and the voiceless plain and pharyngealized stops do not differ in the duration of the aspiration.

Fricatives

A method was developed to describe fricatives acoustically by measuring the spectral center of gravity and the dispersion of spectral energy on the frequency scale from critical band spectra. The fricatives were then plotted in this "space". In addition spectral intensity was considered. See Norlin (1983) and Svantesson (1983). Figure 2 shows the Chinese and Arabic fricatives

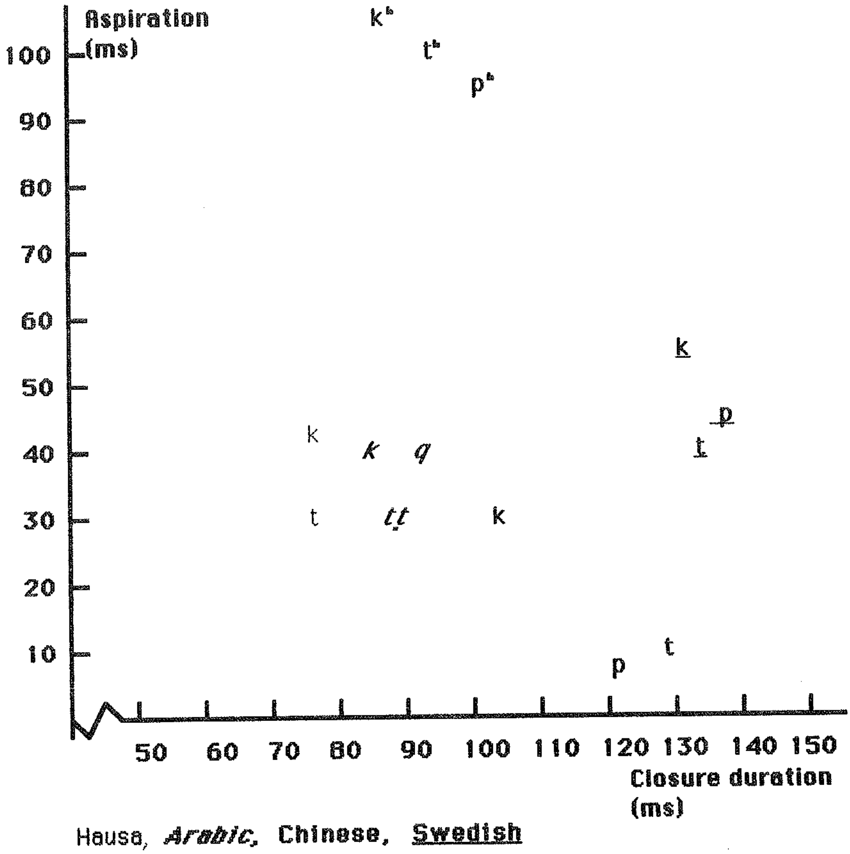


Figure 1. Aspiration and closure duration.

---

plotted with the spectral dispersion against the center of gravity. The fricatives are distributed in two classes, one front and one back. The front fricatives are separated from each other by a combination of these two parameters. The back fricatives of Arabic all have a relatively even distribution of energy with a low center of gravity. It is clear that these two parameters will not separate the back fricatives from each other. An additional parameter of spectral intensity is needed to separate all the fricatives from each other in Arabic.

Comparing the fricatives in the two languages we see, not surprisingly, that the seven voiceless fricatives of Arabic take up a larger part of the available fricative space than the five voiceless fricatives of Chinese. The /ʃ/ fricatives

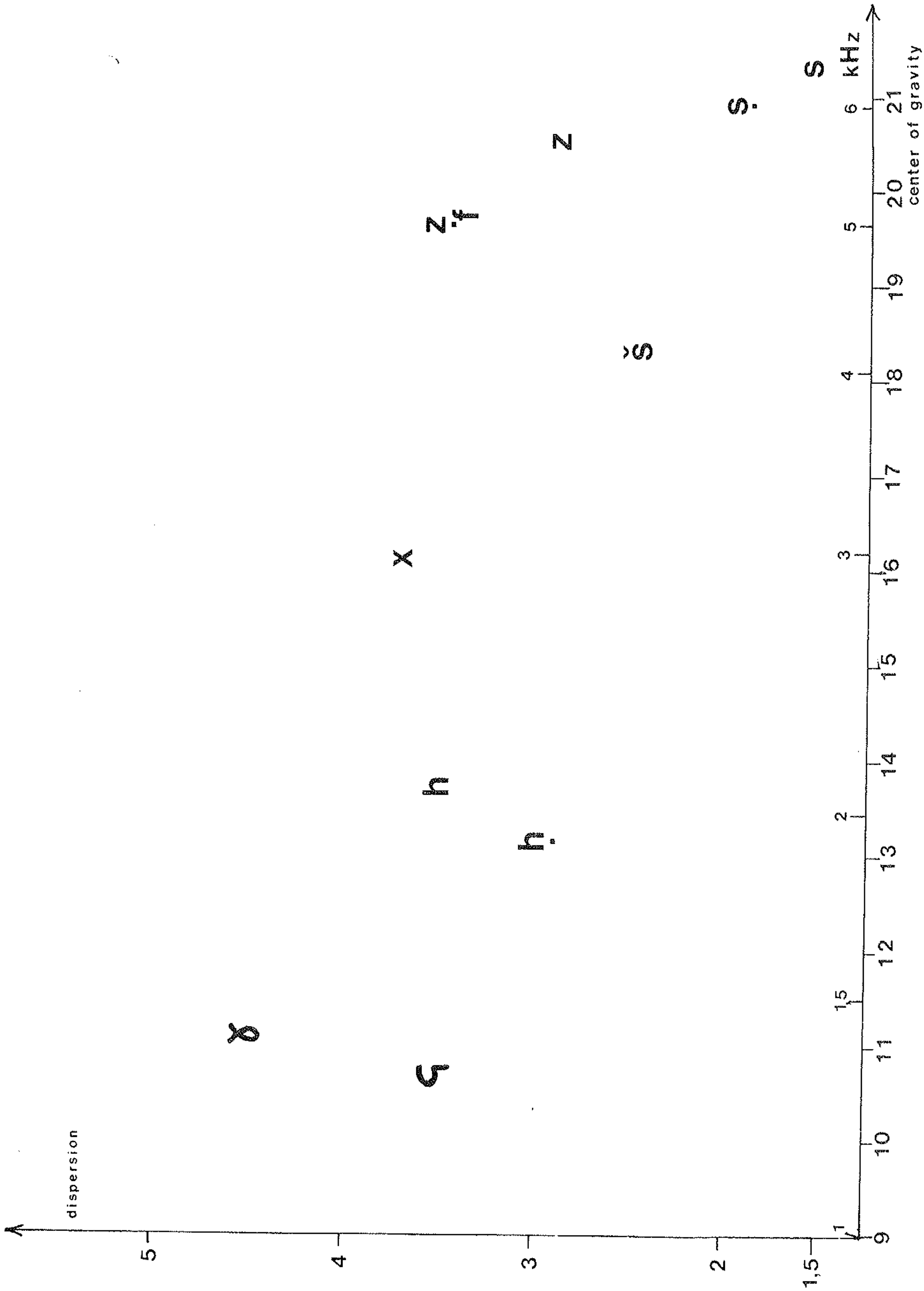


Figure 2a. Arabic fricatives.

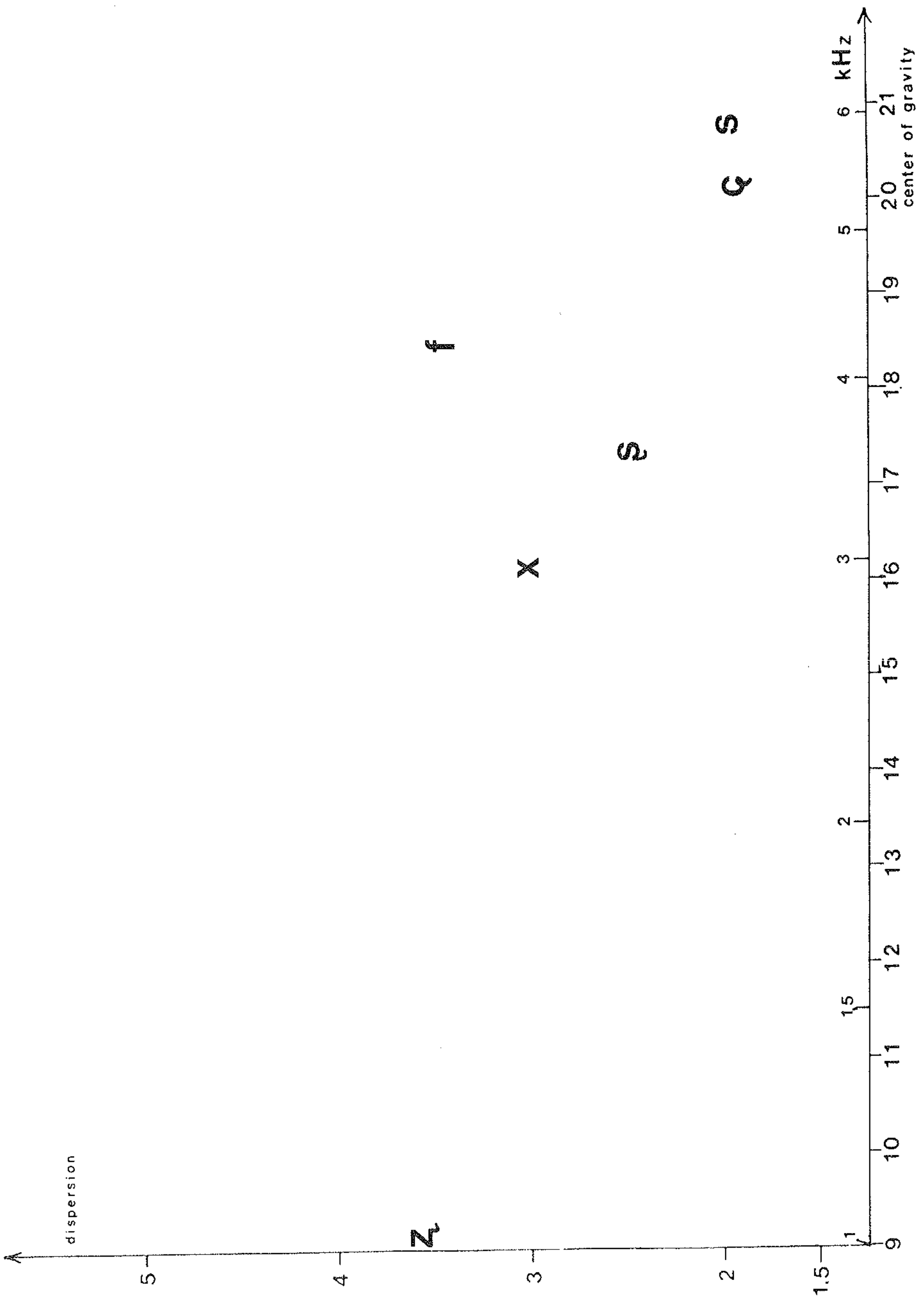


Figure 2b. Chinese fricatives.

in the two languages are very similar, while the /s/ fricatives differ. The Chinese /s/ has a lower center of gravity than the Arabic /s/.

Vowels

Hausa and Arabic are examples of languages with the most common vowel system of all, a five vowel system with the vowels distributed more or less along the periphery of the acoustic vowel space. Both languages have five long and three short vowels. But in spite of their common phonological system and their genetic relationship the vowel systems of Hausa and Arabic do not behave in the same way.

Acoustic properties of Hausa vowels and diphthongs were investigated. The results show that Hausa is best described as having a five vowel system, where these five basic vowels have the qualities of the long vowels. Long vowels are derived as double basic vowels. Phonetically the long vowels are about twice as long as the short ones. The qualities of the short vowels are significantly different from those of the long vowels, but these quality differences can be accounted for by an undershoot mechanism in the speech production.

The results also indicate that the vowel system is currently undergoing changes. Figure 3 shows a formant chart of the long vowels in Hausa from the same environment of between alveolar consonants. The long /oo/ has merged with long /uu/. The formant frequencies of these two vowels are not significantly different. In other environments, the /oo/ is still somewhat lower than /uu/, so the merging of /uu/ and /oo/ is not complete. The basic vowel system may also be additionally modified by the fact that the diphthong /ai/ in most environments has lost its diphthongal quality and monophthongized to long [ee]. However, this monophthongized long [ee] is not the same as the basic long /ee/. Figure 4 shows a formant chart of the monophthongized long [ee] in comparison with the basic long /ii/ and /ee/. Although there is some overlap between the two long [ee] vowels, these two vowels are nonetheless significantly different (paired t-test: $p<0.005$).

Thus the long vowels in Hausa seem to be in the process of transition from a common type of symmetric five-vowel system to an asymmetric system of /ii/, [ee]>/ai/, /aa/, /uu/.

Figure 5 is a formant chart of the three short vowels in Hausa. Both charts show variation between speakers for each vowel, but the back vowels vary considerably more than the front vowels. The tendency for more variation in the back vowels has also been demonstrated for other languages (Keating and Huffman1984).

In Arabic, long plain vowels form well separated clusters with some overlapping for /ii/ and /ee/. Short plain vowels are central compared with the long ones, except short /a/ which occupies approximately the same place as long /aa/. Pharyngealized long and short vowels show the same relationships as the plain ones, except for some overlapping of /uu/ and /oo/. Comparison of the

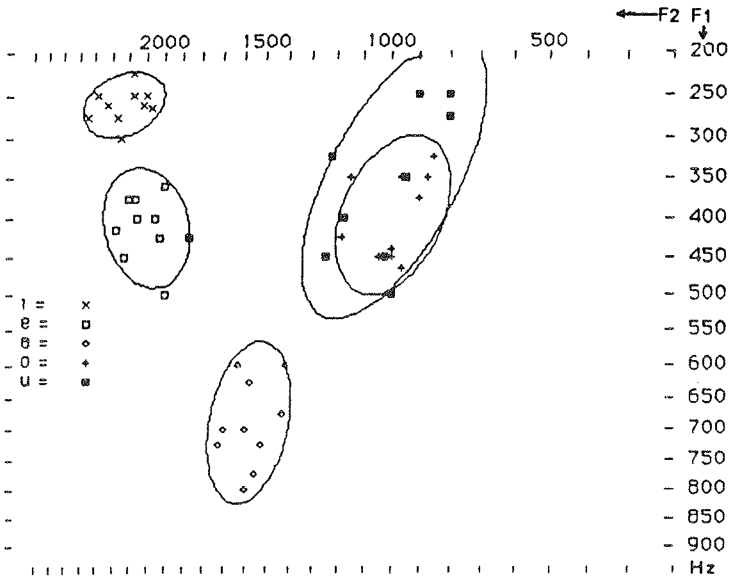


Figure 3. Long vowels in Hausa.

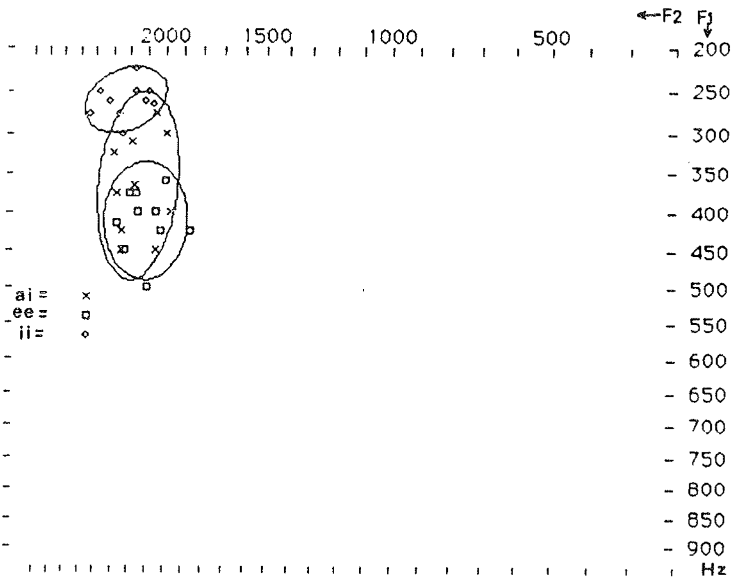


Figure 4. Hausa [ee].

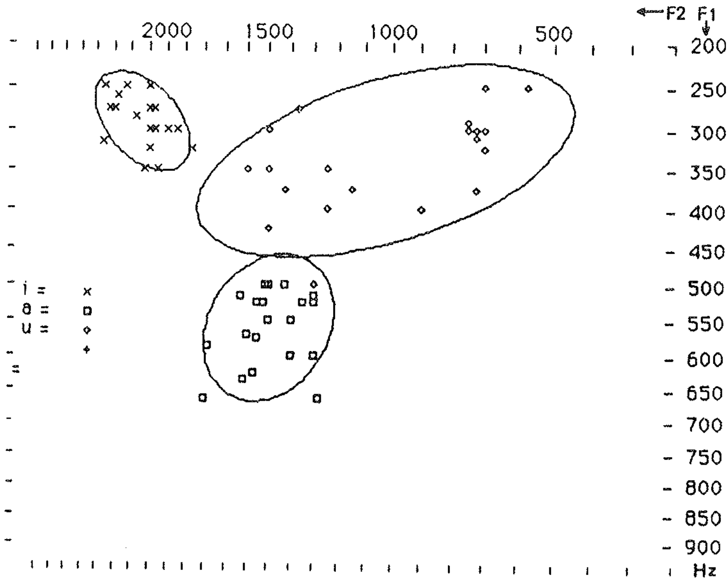


Figure 5. Short vowels in Hausa.

---

vowels in plain and pharyngealized contexts shows differences related to the features high-low and front-back. The difference between plain and pharyngealized allophones is highly significant for long /aa/, is less prominent for /ii/ and /ee/, whereas there is only a small difference or no difference at all for /uu/ and /oo/ (Figure 6). On the other hand, short pharyngealized vowels are more back than short plain ones (Figure 7).

The Chinese vowel system is an example of an unusual distribution of vowels. Chinese has the peripheral /i/, /u/, /a/ and /y/, as well as a middle vowel /ɤ/ with the allophones [ɤ], [o] and [e]. Figure 8 shows the formant chart of Chinese. All the vowels show about the same amount of variation between speakers. The vowels /i/ and /y/ overlap on the F1 - F2 chart, but when the third formant is considered, there is a significant difference in F3 between these two vowels.

Diphthongs

The diphthongs /ai/ and /au/ were studied for Hausa, Arabic, Chinese, and English (Lindau-Webb et al. 1985). Data from American English (Gay 1968) was also used for comparison. A model was devised where the diphthongs are described in terms of formant frequencies of steady state vowels linked by a transition. The interpolation is described using a trinomial equation.

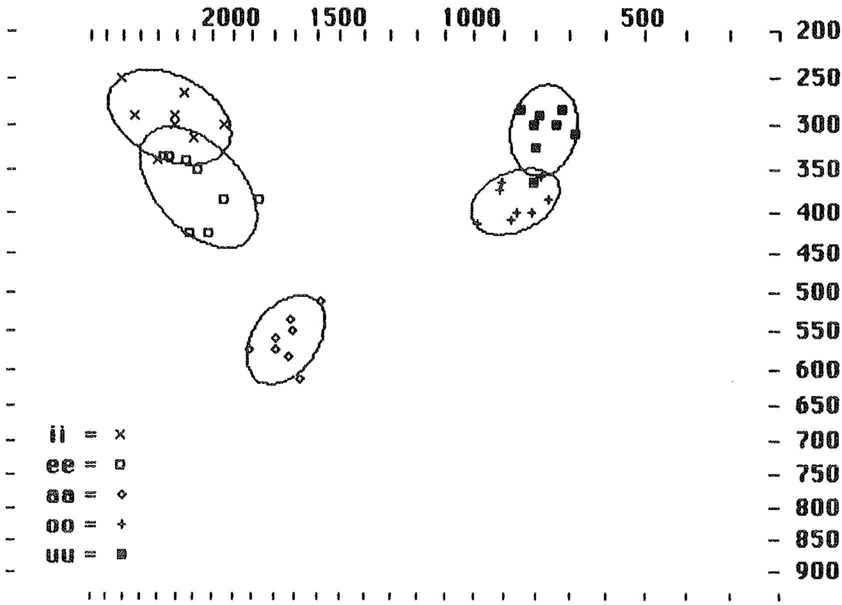


Figure 6a. Long plain vowels in Arabic.

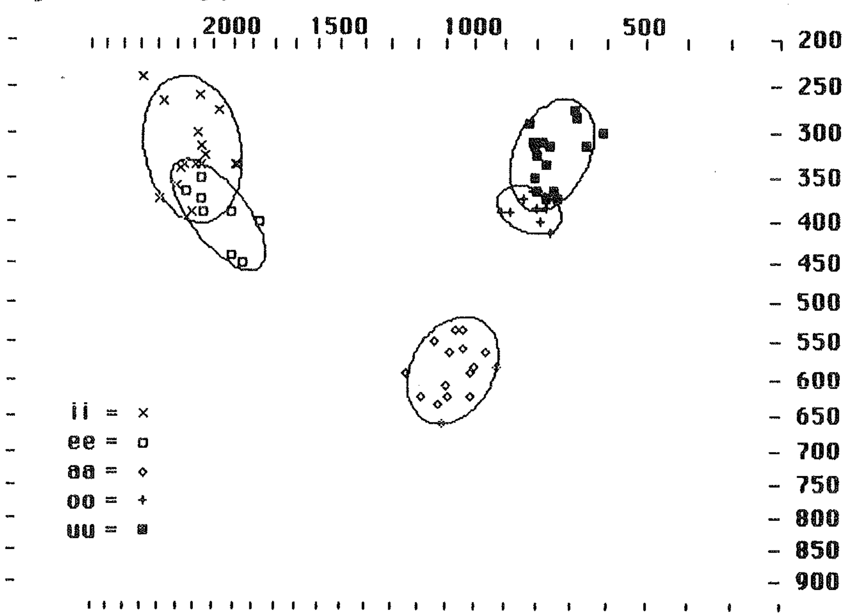


Figure 6b. Long pharyngealized vowels in Arabic.

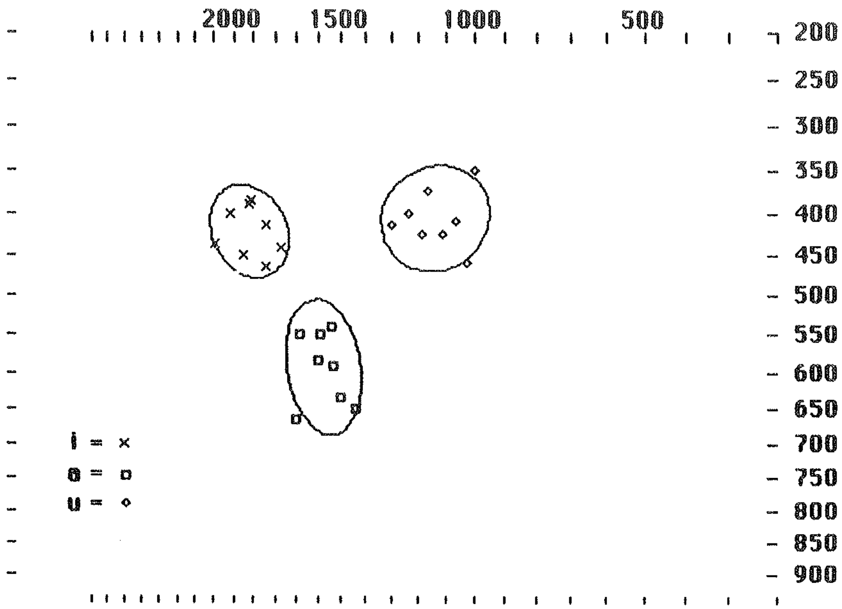


Figure 7a. Short plain vowels in Arabic.

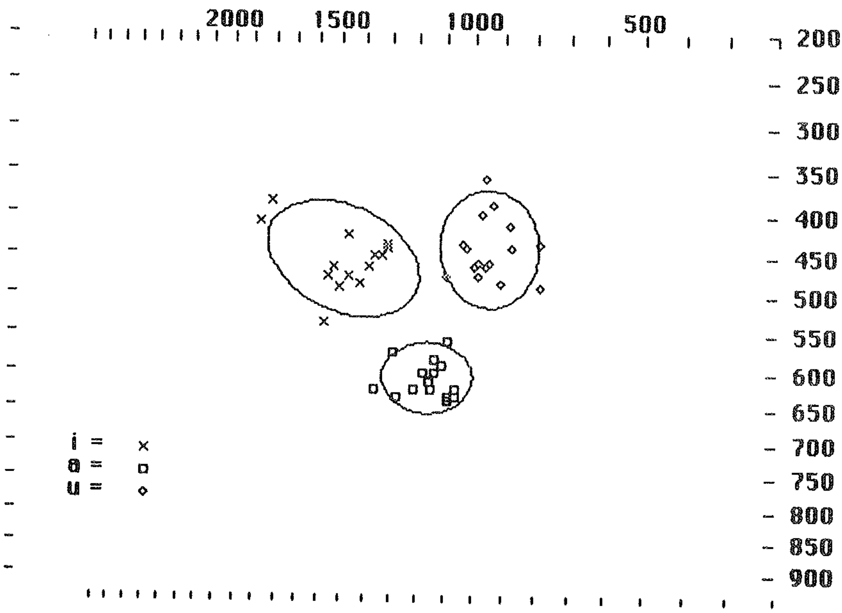


Figure 7b. Short pharyngealized vowels in Arabic.

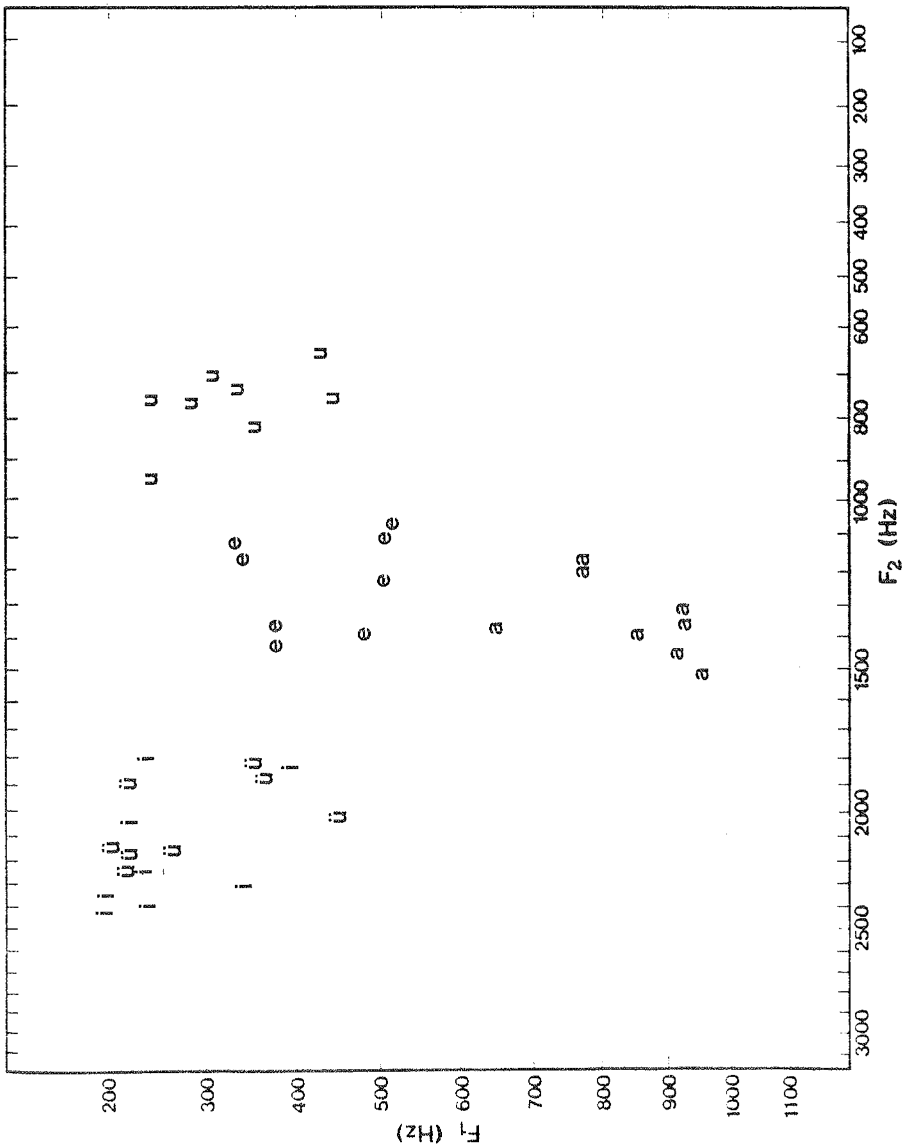


Figure 8. Chinese vowels.

The results show that not only do these diphthongs behave differently in different languages, but the two diphtongs may behave differently from each other within one language, thus supporting language-specific, and even diphthong-specific treatment of diphthongs. In Hausa and Arabic the transition takes up a small percentage of the whole diphthong, while in Chinese and English the tendency is for the transition to take up a large part of the diphthong. Thus the timing of the diphthongal transition is not constant for the "same" diphthong in different languages. The different timing relationships can be predicted from a principle of "the further to go, the longer it takes" (Fischer-Jørgensen 1964) for the /ai/ diphthong, but not for /au/. Thus the transitional rate and duration may be language specific, and even diphthong specific. Also the strategy of the transition taking longer time, if it has a longer time to go is not universal. Later research at UCLA demonstrated that in California English and in Japanese, speakers tend to follow the opposite strategy of "the further the faster".

Prosody

A similar material consisting of statements and questions in focus-free sentences and statements with focus in one of three possible positions was collected for all three languages (examples of pitch curves from the three languages are given in Figure 9). The descriptive framework developed at the Phonetics Department of Lund University over a number of years was used in the analysis.

The assumptions behind this model have recently been summed up in the following way (Gårding 1985):

1. Global intonation stretching over a phrase or a sentence can be separated from local intonation bearing on lexical accents and tones by regarding the accents and tones as superimposed on the global intonation.

2. Any undulating curve (e.g. an intonation curve) can be efficiently described by interpolation between local maxima and minima that we call turning-points.

3. Some of the turning-points for an intonation curve have a rather fixed position relative to the acoustic segments.

4. Giving the positions of these fixed turning-points in time and frequency is an economic way of describing an intonation curve.

5. The local up-and-down structure of an intonation curve usually repeats itself in a global up-and-down structure. This larger pattern is expressed by the tonal grid which in the ideal case is obtained by joining consecutive maxima and minima separately.

6. That part of the grid where the direction or width of the grid is changed or where the grid takes a jump is called a pivot. It marks focus or boundaries of prosodic phrases.

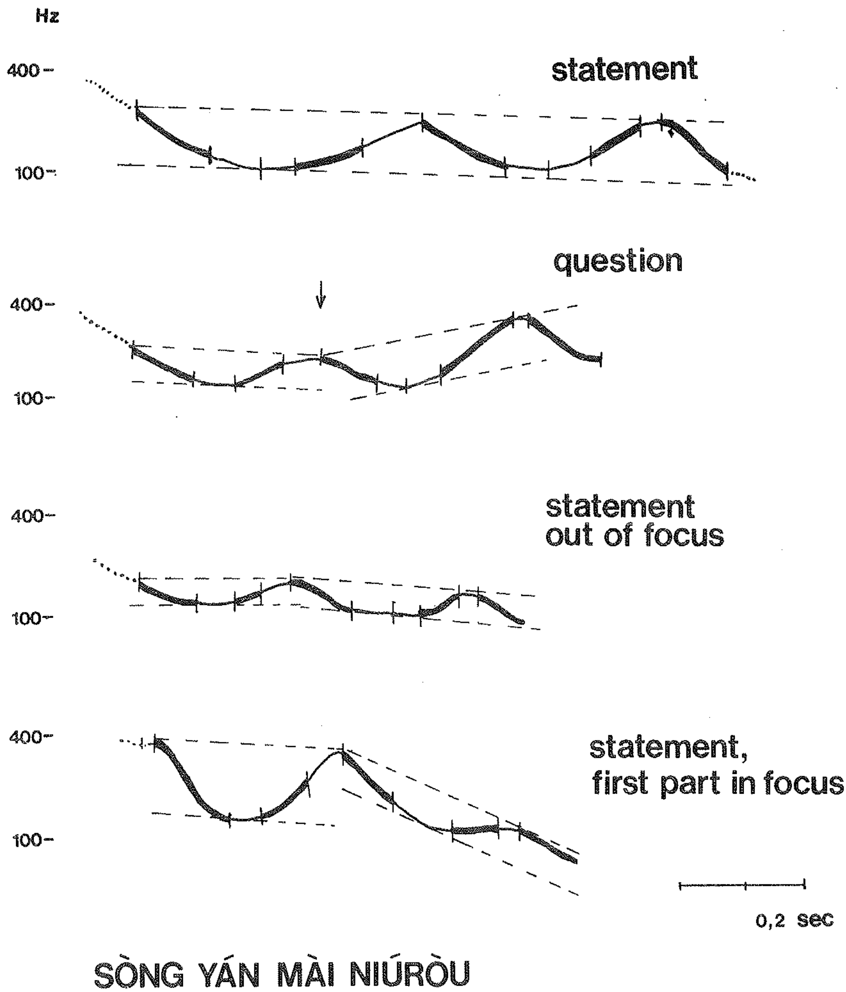


Figure 9a. Intonation curves for Chinese.

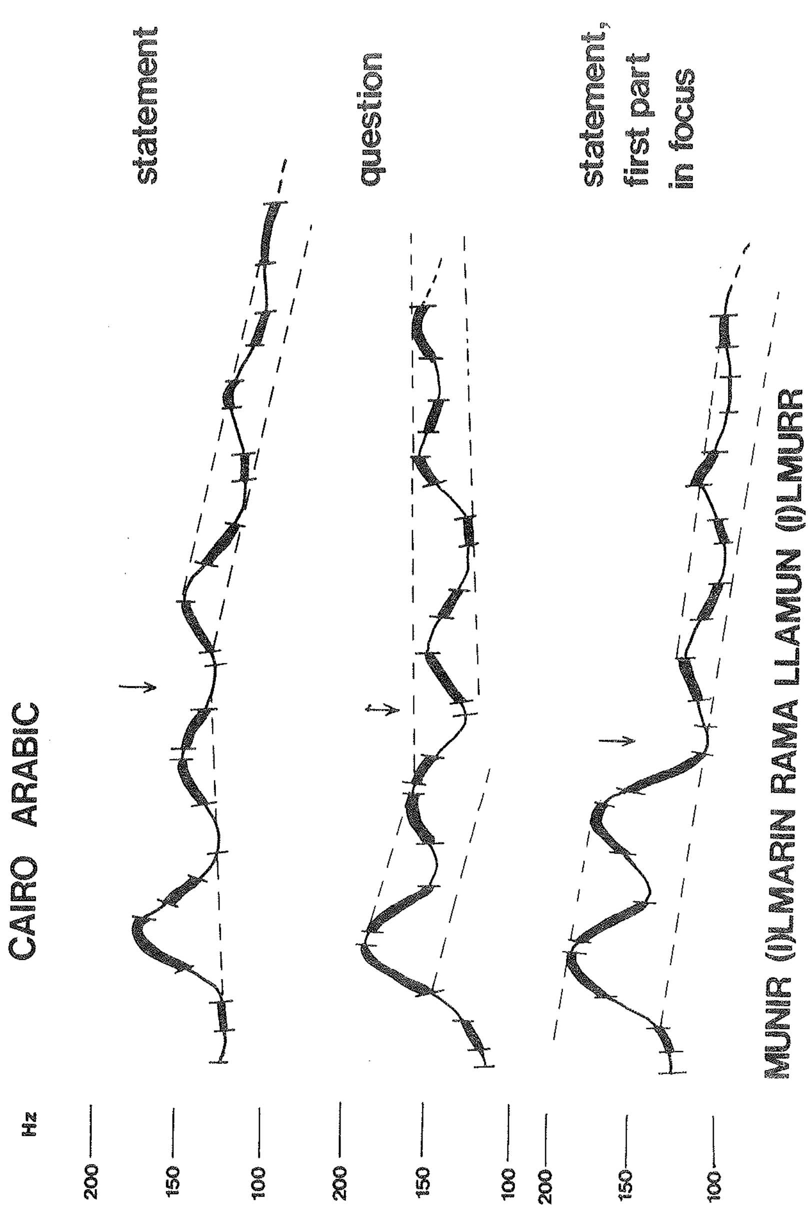


Figure 9b. Intonation curves for Arabic.

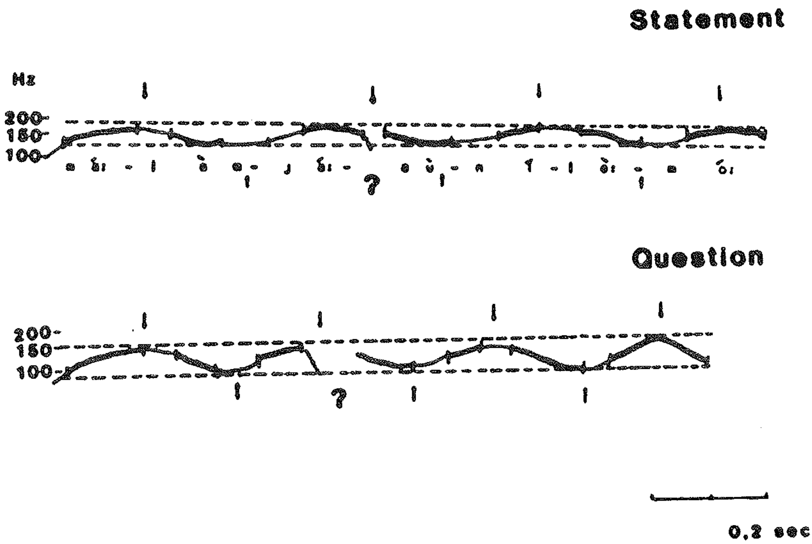


Figure 9c. Intonation curves for Hausa.

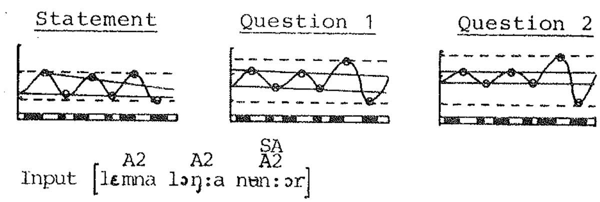
The acoustic correlates behind the concepts of the model and their communicative functions are summarized in Figure 10.

This model of intonation was successfully applied to all three languages, which represent different prosodic systems. Chinese is a tone language with four tones, Hausa is a tone language with two tones, and Arabic is a stress language.

The main result is that once the lexical tones and accents have been factored out, the intonational patterns associated with the global functions of intonation (see Figure 10) are similar.

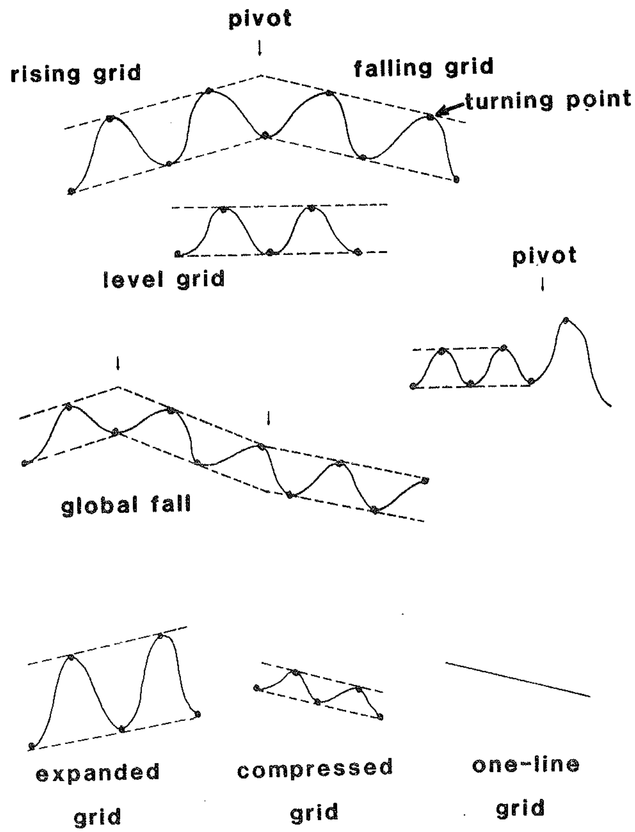
A generative scheme had been suggested for Swedish dialects, in which the global features were to be generated first as a frame of auxiliary lines (later called the tonal grid). The slope of these lines was dependent on speech act type and the length of the utterance. Later the accents were inserted in relation to the grid and the segments according to lexical assignment rules typical of a particular dialect. After some context adjustments the final curve could be obtained by interpolation between the points (Bruce and Gårding 1978).

This scheme had also been applied to syntactically marked and unmarked questions which had been analyzed as having a dialect-independent frame of more or less straightened out and narrowed auxiliary lines expressing question intonation with the same dialect-independent timing of the accentual turning-points as in the statement (Gårding 1979, p. 213):



The difference between the intonation types is thus a difference in the grid, whereas the rules that generate the highs and lows pertaining to the accents remain the same, except for an optional local rule producing a terminal rise in questions (Gårding 1983, p. 21).

This model was applied to Chinese in a qualitative way by Gårding, Zhang and Svantesson (1983). Interspeaker and intraspeaker variability in the tonal (F_0) signal was studied in the Chinese material. The results show the usefulness of the acoustic parameters of the model, turning-point, tonal grid and pivot, which permit rather precise statements. The main observations are that the four speakers use the same lexical and intonational patterns with turning-points very much fixed relative to the segments. From this follows that certain falls and rises are also fixed. Pivots (correlated with phrases) differ from speaker



Intonation parameters	Function Semantic	Syntactic
turning points	words, morphemes	d:o
pivots	constituents (theme/rheme)	d:o (subject/predicate)
grid:direction	speech act type	sentence type
grid width, position	information weight (focus)	clause type

Figure 10. Concepts of the model and their communicative functions.

to speaker, showing their dependence on tempo and style. When pivots do occur, they are in the same location correlated to the syntactic/semantic structure. There is also variability in the manifestation of sandhi rules, which can be regarded as signals of semantic/syntactic coherence. Individual variation was found in the voice range, which varied from one speaker's declamatory style of two octaves to one octave used by the others (Gårding 1985).

Lindau (1986) developed an algorithm that generates schematic F_0 curves of simple statements of two different lengths and questions in Hausa following the general principles of the intonation model above: Rules for intonation and tones are separated and intonation is represented as grids of (near) parallel lines, inside which tones are placed. The direction of the grid lines is associated with sentence type, with a downward slope for statement and straightened out lines for question. The tones are associated with turning-points of the F_0-contour. These turning-points tend to have fixed locations at the end of the syllable with the associated tone. A high tone has a high turning-point in the grid and a low tone has a low one.

Local rules may also modify the exact vertical placement of a tone within the grid. The continuous F_0 contour is modeled by concatenating the tonal points using polynomial equations. Thus the final pitch contour is modeled as an interaction between global and local factors. As for Swedish and many other languages the slope of the intonational grid was found to depend on the type of sentence (statement or question), and the length of the sentence. In addition, the slope of the grid in Hausa is also affected by the tone pattern of the sentence.

The data demonstrate clearly the independence of global and local factors. An observation strengthening this view is that the intonation of sentences consisting of high tones only exhibits a downwards slope. This kind of slope cannot be explained by reference to local rules of downstepping but is best described as a manifestation of global intonation.

As in the other languages studied in the project, the different speakers made use of very different pitch ranges in sentences on alternating High and Low tones. This was attributed to non-linguistic factors of attitude and personality of the speaker. As has been shown for the other languages, more involvement and interest results in a larger pitch-range (see also Bruce 1982).

In addition to the straightened out grid lines in questions, the last High tone of the sentence is locally raised, sometimes followed by a fall. An interesting feature is that this raised High also tends to have a delayed turning-point. This delay is in Hausa related to the height of the peak, conforming to the same principle that we found for the diphthongal transitions: the further to go, the longer it takes.

Question-word questions are characterized by a slight downward slope, where the amount of slope is something in between that of statements and yes-no questions. These results are similar to those of Gårding (1979) and Thorsen (1978). Thorsen concluded that the more morphosyntactic cues there were in the questions, the more the slope looked liked that of statements. Question intonation in Arabic and Chinese also follows this pattern. In other words, the less morphosyntactic cues to its type the sentence carries, the more work will the intonation have to do.

Perception

Chinese, with its four tones, is a suitable testing ground for an intonation model. While in Sweden, Zhang, using the ILS program, explored the importance of the timing of turning-points in relation to the grid for the recognition of specific tones. This work was continued by Gårding, Kratochvil and Svantesson, who could list features which appeared to be invariant for a particular tone as a result of perceptual tests of synthetic stimuli which had configurations intermediary between Tone 3 and Tone 4. One invariant feature was that the first half of the tone (including the fall) seemed to be important for Tone 4 and the second part for Tone 3. It seemed, then, that the significant parts of the tones both included changes from one mode to another mode (i.e. turning-points). The result may not be without general perceptual significance.

Effects of the project

For Chinese scholars, our intonation model is attractive because it offers a simple way of handling the interaction between tone and intonation. Collaboration has continued and is continuing between Lund and the Acoustics Institute of the Academia Sinica, Beijing. Shi Bo, a graduate student from that institute, who is studying in Lund with a scholarship from the Swedish Institute, is concerned with the perception and identification of global features, as part of her intention of making the model more quantitative. Her stimuli were tested during Gårding's 1986 visit to the Academia Sinica. The perception of tonal movements is also interesting to David House who is trying to give the model a perceptual orientation (see also Gårding and House 1985). Our work on diphthongs is now being continued at UCLA.

Pedagogical applications

The pedagogical applications of our project have not been completed. A manual with pronunciation drills for Chinese with phonetic illustrations is now being completed by Magnus Nordenhake. This book is a Swedish translation and adaptation of a manual for English students by Paul Kratochvil. Taghrid Anbar and Kjell Norlin are collaborating on a contrastive study of Arabic and Swedish.

Summary

The project has resulted in new material from languages which are given preferential treatment by Lund University. The segmental systems, in particular stops, fricatives and vowels have been thoroughly investigated. The same methods, some of them developed within the project, were used for all three languages, thus enabling direct comparison between them and laying a foundation for future research on other languages and for pedagogical applications.

Prosodic features have been successfully analyzed in all three languages by means of the intonation model developed here. The model has been implemented in an algorithm producing statements and questions in Hausa. Apart from this technical achievement, the model serves as a convenient frame for further explorations of the interaction of tone (accent) and intonation. At the same time we have strengthened our claim that the model gives a general frame for any prosodic system and that it sheds light on the structure of intonation in general.

References

(Publications from the LUCLA Project are given in the Bibliography, below)

Bruce, G. 1982. "Developing the Swedish intonation model". Working Papers (Dept. of Linguistics, Lund) 22, 51-116.

Bruce, G. and E. Gårding. 1978. "A prosodic typology for Swedish dialects". In E. Gårding, G. Bruce and R. Bannert, eds. Nordic Prosody, Papers from a Symposium, 219-28. Lund, Dept. of Linguistics.

Fischer-Jørgensen, E. 1964. "Sound duration and place of articulation". Zeitschrift für Phonetik und allgemeine Sprachwissenschaft 17, 175-207.

Gårding, E. 1979. "Sentence intonation in Swedish". Phonetica 36, 207-15.

Gårding, E. 1983. "A generative model of intonation". In Cutler and Ladd, eds. Prosody: Models and measurements, 11-25. Springer.

Gårding, E. and D. House. 1985. "Frasintonation, särskilt i svenska". In Svenskans beskrivning 15, 205-21.

Gay, T. 1968. "Effect of speaking rate on diphthong formant". The Journal of the Acoustical Society of America 44:6, 1570-75.

Keating, P. and M. Huffman. 1984. "Vowel variation in Japanese". Phonetica 41, 191-207.

Löfqvist, A. 1976. "Closure duration and aspiration for Swedish stops". Working Papers (Dept. of Linguistics, Lund) 13, 1-40.

Thorsen, N. 1978. "Aspects of Danish intonation". In E. Gårding, G. Bruce and R. Bannert, eds. Nordic Prosody, Papers from a Symposium, 23-32. Lund, Dept. of Linguistics.

Appendix

Consonant systems:

Hausa:

	b	t	d	c	ɟ	k	g	k^j	g^j	k^w	g^w	ʔ
	ɓ		ɗ			k'		k^j'		k^w'		
ɸ		s	z	ʃ								h
		s'										
	m		n									
			l									
			r									
			ɽ									
	w				j							
					j̰							

Arabic:

	b	t	d			k	g	q		ʔ
		ṭ	ḍ							
f		s	z	ʃ		χ	ʁ	ħ	ʕ	h
		ṣ	ẓ							
	m		n							
			l							
			r							
	w				j					

Chinese:

p^h	t^h						k^h	
p		t						k
		ts^h		$tʂ^h$		$tɕ^h$		
		ts		tʂ		tɕ		
f		s		ʂ	ʐ	ɕ		x
	m		n		ŋ			
					l			

Bibliography of the LUCLA project

Gårding, Eva. 1984. "Comparing intonation". Working Papers (Dept. of Linguistics, Lund) 27, 75-99.

Gårding, Eva. 1984. "Chinese and Swedish in a generative model of intonation". In C.-C. Elert, I. Johansson, and E. Strangert, eds. Nordic Prosody III, 79-91.

Gårding, Eva. 1985. "In defence of a phrase-based model of intonation". Working Papers (Dept. of Linguistics, Lund) 28, 1-18.

Gårding, Eva. 1985. "Constancy and variation in Standard Chinese tonal patterns". Working Papers (Dept. of Linguistics, Lund) 28, 19-51.

Gårding, Eva. 1985. "En intonationsmodell och dess förankringar". In: E Ahlsén, J Allwood and E. Hjelmquist, ed. Tal-Ljud-Hörsel 2 (Guling 13), 86-102. Göteborg.

Gårding, Eva. 1985. "Superposition as an invariant principle of intonation". In J. Perkell and D.H. Klatt, eds. Invariance and variability in speech, Hillsdale, NJ: L Erlbaum Assoc. Publ.

Gårding, Eva, Paul Kratochvil, Jan-Olof Svantesson and Zhang Jialu. 1985. "Tone 4 and Tone 3 discrimination in Modern Standard Chinese". Working Papers (Dept. of Linguistics, Lund) 28, 53-67.

Gårding, Eva, Zhang Jialu and Jan-Olof Svantesson. 1983. "A generative model for tone and intonation in Standard Chinese". Working Papers (Dept. of Linguistics, Lund) 25, 53-65.

Lindau, Mona. 1983. "Glottalic consonants in Hausa and some other languages". Working Papers (Dept. of Linguistics, Lund) 25, 85-98.

Lindau, Mona. 1984. "Testing a model for Hausa intonation". Working Papers (Dept. of Linguistics, Lund) 27, 145-63.

Lindau, Mona. 1984. "Phonetic differences in glottalic consonants". Journal of Phonetics 12:2, 147-57.

Lindau, Mona. 1985. "The story of /r/". Ch. 11 in Victoria Fromkin, ed., Linguistic phonetics. Essays in honor of Peter Ladefoged, 157-68. Orlando: Academic Press.

Lindau, Mona. 1985. "Hausa vowels and diphthongs". UCLA Working papers in phonetics 60, 40-54.

Lindau. Mona. 1985. "Testing a model of intonation in a tone language". UCLA Working papers in phonetics 61, 26-39.

Lindau, Mona. 1985. "Hausa vowels and diphthongs". Studies in African Linguistics 16:2, 161-182.

Lindau. Mona. 1986. "Testing a model of intonation in a tone language". The Journal of the Acoustical Society of America. In press.

Lindau, Mona and Peter Ladefoged. 1985. "Variability of feature specifications". Ch. 21 in J. Perkell and D.H. Klatt, eds. Invariance and variability in speech, 464-79. Hillsdale, NJ: L Erlbaum Assoc. Publ.

Lindau, Mona, Kjell Norlin and Jan-Olof Svantesson. 1985. "Cross-linguistic differences in diphthongs". UCLA Working papers in phonetics 61, 40-44.

Nordenhake, Magnus and Jan-Olof Svantesson. 1983. "Duration of Standard Chinese word tones in different sentence environments". Working Papers (Dept. of Linguistics, Lund) 25, 105-111.

Norlin, Kjell. 1983. "Acoustic analysis of fricatives in Cairo Arabic". Working Papers (Dept. of Linguistics, Lund) 25, 113-137.
Norlin, Kjell. 1984. "Acoustic analysis of vowels and diphthongs in Cairo Arabic". Working Papers (Dept. of Linguistics, Lund) 27, 185-208.
Norlin, Kjell. 1985. "Acoustic analysis of vowels and diphthongs in Cairo Arabic". In R Schuh, ed. Précis from the 15th Conference on African Linguistics (Studies in African Linguistics, Supplement 9), 238-44. Los Angeles: UCLA.
Norlin, Kjell. 1985. "Pharyngealization in Cairo Arabic". Working Papers (Dept. of Linguistics, Lund) 28, 139-50.
Svantesson, Jan-Olof. 1983. "Acoustic analysis of Chinese fricatives and affricates". Working Papers (Dept. of Linguistics, Lund) 25, 195-211.
Svantesson, Jan-Olof. 1984. "Vowels and diphthongs in Standard Chinese". Working Papers (Dept. of Linguistics, Lund) 27 (1984), 209-35.
Svantesson, Jan-Olof. 1986. "Acoustic analysis of Chinese fricatives and affricates". Journal of Chinese Linguistics (Berkeley) 14, 53-70.
Zhang Jialu. 1984. "Perceptual studies in Chinese intonation using foreign listeners". Working Papers (Dept. of Linguistics, Lund) 27, 237-45.

Presentations

Gårding, Eva, Zhang Jialu and Jan-Olof Svantesson: Tone and intonation in Standard Chinese. Prosody, normal and abnormal, Association psycholinguistique de l'Europe, Zürich, March 1983.

Norlin, Kjell: Fricatives in Cairo Arabic. Finska orientsällskapet, Helsingfors, March 1983.

Lindau, Mona: Question intonation in Hausa. 14th Conference on African Linguistics, Madison, April 1983.

Gårding, Eva: Chinese and Swedish in a generative model of intonation. Nordic Prosody III, Umeå, June 1983.

Gårding, Eva: A generative model for tone and intonation in Standard Chinese. Institut für Phonetik und sprachliche Kommunikation, München, June 1983.

Zhang Jialu: Chinese tones and intonation according to a Swedish model. Academia Sinica, Beijing, July 1983.

Gårding, Eva: Intonation units and pivots, syntax and semantics. Contribution to symposium, 10th International Conference on Phonetic Sciences, Utrecht, August 1983.

Gårding Eva, Jan-Olof Svantesson and Zhang Jialu: Tone and intonation in Standard Chinese. Tenth International Congress of Phonetic Sciences, Utrecht, August 1983. Abstract in: A. Cohen and M.P.R. van den Broecke, ed. Abstracts of the Tenth International Congress of Phonetic Sciences, p. 605. Dordrecht: Foris, 1983.

Lindau, Mona: Intonation in Hausa. Tenth International Congress of Phonetic Sciences, Utrecht, August 1983. Abstract in: A. Cohen and M.P.R. van den Broecke, ed. Abstracts of the Tenth International Congress of Phonetic Sciences, p. 611. Dordrecht: Foris, 1983.

Gårding, Eva: Superposition as an invariant feature of intonation. Symposium on Invariance and Variability, MIT, October 1983.

Gårding, Eva: Superposition är en invariant princip i intonation. Institute of Linguistics, Stockholm, November 1983.
Lindau, Mona: Vowels and diphthongs in Hausa. 15th Conference on African Linguistics, Los Angeles, March 1984.
Norlin, Kjell: Acoustic analysis of fricatives and vowels in Cairo Arabic. 15th Conference on African Linguistics, Los Angeles, March 1984.
Svantesson, Jan-Olof: Some results from the LUCLA project. Institute of Acoustics, Academia Sinica, Beijing, Sept 1984.
Norlin, Kjell: Arabisk fonologi. Humanistdagarna i Lund, March 1985.
Gårding, Eva: En intonationsmodell och dess förankringar. Tal-Ljud-Hörsel, Göteborg, March 1985.
Lindau, Mona, Kjell Norlin and Jan-Olof Svantesson: Resultat från projektet LUCLA. Tal-Ljud-Hörsel, Göteborg, March 1985.
Norlin, Kjell: Faryngalisering i arabiska. Tal-Ljud-Hörsel, Göteborg, March 1985. Abstract in: E Ahlsén, J Allwood and E. Hjelmquist, ed. Tal-Ljud-Hörsel 2 (Guling 13), p. 177. Göteborg.
Lindau, Mona, Kjell Norlin and Jan-Olof Svantesson: Cross-linguistic differences in diphthongs. 109th Meeting of the Acoustical Society of America, Austin, Texas, April 1985. Abstract in: The Journal of the Acoustical Society of America, 77, Supplement 1, p. 100.
Gårding, Eva: Prosodic constancy and variation across languages. Miniconference on phonology and phonetics, Institute of Linguistics, Stockholm, May 1985.
Norlin, Kjell: Pharyngealization in Arabic. Miniconference on phonology and phonetics, Institute of Linguistics, Stockholm, May 1985.
Gårding, Eva: En intonationsmodell och dess förankringar. Helsingfors and Åbo, May 1985.
Gårding, Eva: An intonation model and its anchors. Aix en Provence, July 1985.
Gårding, Eva: A phrase based model of intonation. München, July 1985.

A simple qualitative model for the vibration of the vocal folds

Lars Gårding

The purpose of this short note is to establish a simple qualitative model for the vibration of the vocal folds. Its basic assumptions are physically credible, its main aim is pedagogical and it can be explained with a minimum of mathematics. It is no real substitute for the more ambitious models of Flanagan and Landgraf (1975) and Tietze (1973) which take into consideration that the closure of the vocal folds is not uniform. The lower part closes first. Therefore these models represent the vocal folds by two masses on each side. Our model just uses one mass. Its assumptions are as follows.

A. The tensions and the masses of the muscles involved, mainly the vocalis and the cricothyroid muscles, determine a damped vibratory system S with a certain frequency $w/2\pi$ Hz.

B. The duration of the closed phase is half of the period of S.

C. The transglottal pressure P and a rebound B from S initiate the opening movement of the system with a certain force Q of very short duration.

D. When the movement initiated by Q has reached its maximum, a Bernoulli force R sets in which decreases proportionally to the opening between the vocal cords.

The assumption A is an uncontroversial simplification. The basic frequency w is determined by the tension T of S and its mass M in such a way that it decreases as T decreases and M increases. (The theoretical formula in the undamped case is $w^2=T/M$). The assumption B is motivated mainly by the observation that the closed and open phases of the glottal cycle are approximately equal. Since the modifications of the frequency w introduced by the Bernoulli force are never very large (see below), it has seemed natural to tie the duration of the closed phase to the state of the system S.

The assumptions C and D express a convenient way of circumventing the complicated interaction between the movements of the vocal cords and the flow through the glottis. D expresses the conventional view that the Bernoulli sucking force helps close the glottis but adds to it the assumption that the force sets in when the opening is maximal. This is natural because the Bernoulli force is largest when the flow is stationary and the natural moment for this to happen is when the opening of the glottis is maximal. As the glottis closes, the flow becomes less and less stationary. The assumption that the decrease of R follows the size of the opening is somewhat ad hoc but not unreasonable.

Note that the assumption C considers the force Q to be composed of subglottal pressure P and a rebound B. Under stable phonation, Q is of course constant. For a beginning phonation, the rebound vanishes for the first cycle but picks up to a stable value later. This fits with the increasing amplitudes of a beginning phonation.

The model seems to be able to explain a number of known facts about the vibration of the vocal folds.

Figure 1 shows the changing shape of the glottal cycle when S,P,Q are fixed and the Bernoulli force R increases /1/. When R is zero, the assumption A says that the open part of the glottis cycle follows a damped sine curve whose maximum is proportional to Q/M. The sine curve in the figure is only slightly damped. We see that the closing of glottis becomes faster as R increases. In this way, the curve representing the open phase acquires its characteristic asymmetric form (see e.g. Fant 1979 and Anathapadmanabha and Fant 1982) at the same time as the period of the glottal cycle decreases, i.e. as the frequency increases.

An increased Bernoulli force can be thought of as an efficient way of producing sound. Favourable conditions for this are regular movements of the vocal cords with no asymmetries or irregularities. It is probable that trained singers realize these conditions. A certain support for this statement is the fact that the ratio of the duration of the open to the closed phase is smaller for trained singers than for untrained ones (Sundberg and Gauffin 1979).

It can be shown mathematically (see the formula (3) of the appendix) that if S is fixed and Q and R are proportional to transglottal pressure P, then the duration of the glottal cycle is independent of P (Fig.2). But this is only approximately true. As a matter of fact, experiments have shown that when transglottal pressure increases, the duration of the glottal cycle decreases so that the frequncy increases. A reasonable explanation for this second order effect is that the opening force Q is not proportional to P for large values but subproportional, i.e. it increases less than proportionally to P. Since the maximal glottis opening is proportional to Q, this means that it grows less than proportionally to P. Then the formula (3) of the appendix shows that this has the same effect as an increase of the Bernoulli force R. Hence the duration of the glottis cycle decreases a little when P increases (Fig. 1).

Another thing that can be shown mathematically in the model is that the effect of the Bernoulli force R decreases when w increases as a result of increased tension and/or decreased mass of the system S. This means that if w is large, there is less asymmetry in the glottis curve, an effect which has been observed in a striking way by the sine-like glottograms of falsetto voice (Sundberg 1980 p. 64) /2/.

Mathematical appendix.

According to A, the glottal opening x as a function of t, $x=x(t)$ satisfies the equation

$$Mx''(t) + CMx'(t) + Tx(t) = 0$$

during the open phase. The second term on the left accounts for the damping. For simplicity in this account, we put $C=0$. This restricts our formulas to the undamped or slightly damped case, but no new effects will appear unless the damping is very large.

Shortly after the impact of transglottal pressure, the function $x(t)$ has the form

(1) $$x(t) = (Q/Mw)\ \sin wt\ ,\ w^2 = T/M,$$

which means that $Mx''(t) + Tx(t) = Q\delta(t)$ (to approximate the short-lived force Q by an instantaneous one seems legitimate). The function $x(t)$ reaches its maximum Q/M when $t=\tau= \pi/2w$. After this, the equation of movement changes to

$$Mx''(t) + Tx(t) + (RM/Q)x(t) = 0.$$

Hence, the equation of the closing phase of the open part of the glottal cycle is

(2) $$x(t) = (Q/M)\ \cos W(t-\tau)$$

with t between τ and $\tau + \pi/2W$ where W is given by

(3) $$W^2 = w^2 + (R/Q).$$

The formula (3) shows how an increased Bernoulli force makes the closure of the glottal cycle steeper and also that the quotient W/w decreases as w increases and R/Q is fixed. It also shows that if R and Q are proportional to P, then the period length $(\pi/w)+(\pi/W)$ is independent of P but that if Q increases less than a constant times P, then R/Q increases with P so that W also increases and hence the period length decreases.

The computations above can only be taken in a qualitative sense, although the model contains enough parameters to permit close fits to very regular observed glottis cycles. In my view, the importance of the model lies in the fact that the parameters have a very direct physical significance and that the model seems to permit direct and meaningful interpretations of many aspects of the glottal cycle.

/1/. The steep beginning of the curve is an artefact due to the assumption that the force Q is instantaneous.

/2/. In falsetto voice, the vocal folds do not close. The mathematical appendix covers this case too, if applied only to the open phases above a mean opening.

REFERENCES

Ananthapadmanabha T.V. and G. Fant (1982): Calculation of true glottal flow and its components. Speech Communication I 167-184, North-Holland.

Fant G. (1979) :Speech production. A glottal source and excitation analysis. STL-QPSR 1/1979, 85-105

Flanagan J.L. and L. Landgraf (1968): Self-oscillating source for vocal tract synthesizers. IEEE Trans. Audio Electroacoust. Vol. AU-16, 1968, 57-64

Sundberg J. (1980) :Röstlära. Proprius, Stockholm.

Sundberg J. and J. Gauffin (1979): Waveform and Spectrum of the Glottal Vaoice Source. Frontiers of Speech Communication Research. Academic Press, 301-320

Tietze I R (1973): The human vocal cords: A mathematical model. Part I, Phonetica 28, 129-170, Part II, Phonetica 29, 1-21

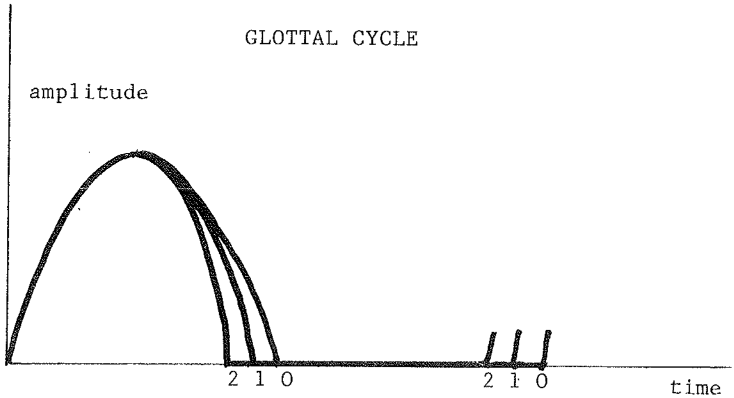


Fig. 1 Asymmetry and shortening of the length of the glottal pulse when the Bernoulli force R increases. The length of the closed phase is constant, curves with the same number correspond.

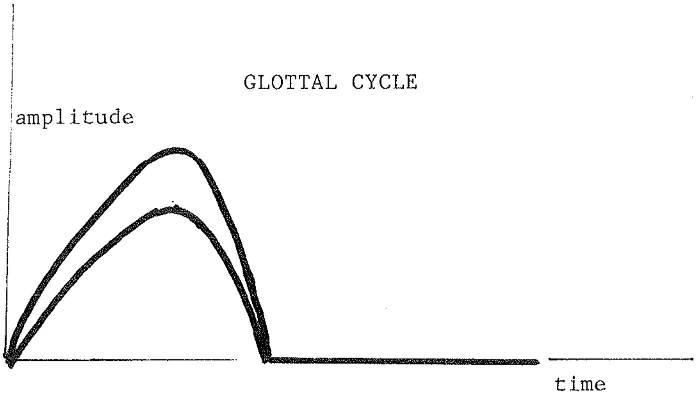


Fig. 2 The length of the glottal pulse is constant when the subglottal pressure P varies and the initial force Q and the Bernoulli force R are proportional to P

Backness and roundness harmony in Hungarian

Magnus Olsson

1. INTRODUCTION

Ferdinand de Saussure said at one of his lectures: "The first thing that strikes us when we study the facts of language is that their succession in time does not exist insofar as the speaker is concerned. He is confronted with a state. That is why the linguist who wishes to understand a state must discard all knowledge of everything that produced it and ignore diachrony."; translation in1960, p. 81.

Alf Nyman - a Swedish philosopher - follows the same line of thought: "But the step of thought from the origin to the value [....] throws itself precipitately between two differing *dimensions* within the world of human judgement: the two standpoints of genetic explanation and estimating reflection.";1960, p. 81 (my translation).

These passages have been of some importance to me when writing this article, because it deals with Hungarian vowel harmony - both the front /back type (hereafter simply called harmony) and the roundness harmony - in *synchronic* terms. Etymological assumptions about certain words and certain phonemes have in previous attempts obscured the real facts concerning vowel harmony. Although historical linguistics is an interesting part of linguistics in its own right and may shed light upon the synchronic study, the two views should not be confused.

Section 2 is a description of the vowel system and the general vowel harmony types in Hungarian. In section 3 we will look at the three main problems concerning harmony. Section 4 consists of formalizations of relevant rules.

2. VOWEL SYSTEM AND VOWEL HARMONY

One way of dividing the Hungarian vowels in subsets is according to backness. Harmonic front and back vowels are normally kept apart wordinternally. Vowels that belong to the neutral group - a subclass of the front vowel class - may appear freely with vowels from any of the two harmonic sets. Harmonic vowels have a much greater influence on the backness of other vowels in the word. Neutral vowels are unchangeable in suffixes while harmonic suffix vowels typically conform to harmony.

(1)

	Short			Long			
	Front		Back	Front		Back	
	-round	+round	+round	-round	+round	-round	+round
High	i	ü	u	í	ű		ú
Mid	ë	ö	o	é	ő		ó
Low	e		a			á	

On the above chart[1] are all the orthographic vowels as well as the phoneme ë, which is spelled <e>. In the seven-vowel dialects (with Budapest as center), ë has merged with e. The previous rule-writers have only been concerned with this dialect, but more than two thirds of the Hungarians in Hungary and neighbouring countries retain ë. Henceforth I will sometimes use forms with ë for explanatory purpose and rules will be given for both standard dialect groups. The relation between pronunciation and spelling is otherwise much closer than e.g. in English, French or Swedish (the main exceptions being certain proper names).

Scholars have different opinions about which vowels are neutral. Vago counts e, é, i and í as neutral, while Ringen´s opinion is that only é, i and í should be termed neutral.

Suffix vowels usually agree in backness with the last root vowel[2]. If that vowel is harmonic or there are only neutral and front harmonic vowels in the stem, this statement is always valid, e.g.:

(2)

		(adess.)	(delat.)	(instr./com.)	(allat.)
well	kút	kútnál	kútról	kúttal	kúthoz
hair	haj	hajnál	hajról	hajjal	hajhoz
ear	fül	fülnél	fülről	füllel	fülhöz
rain	eső	esőnél	esőről	esővel	esőhöz
thorn	tövis	tövisnél	tövisről	tövissel	tövishez
fairy	tündér	tündérnél	tündérről	tündérrel	tündérhez

If the last vowel is neutral and the first harmonic vowel to the left in the morpheme is back, there are three possibilities. Normally the suffix vowels become back, e.g.:

(3)

		(delat.)	(instr./comitat.)
party, spree; rebellion	muri	muriról	murival
eraser	radír	radírról	radírral
coffee	kávé	kávéról	kávéval

But after some roots the suffix vowels are always front[3]:

(4)

		(delative)	(instr./comitat.)
concert	koncert	koncertről	koncerttel
bronchitis	bronchitisz	bronchitiszről	bronchitisszel

The last case is free variation between front and back vocalism:

(5)

		(delative)	(instr./comitat.)
positive	pozitív	pozitívról/ pozitívről	pozitívvel/ pozitívval
fool, greenhorn	balek	balekről/ balekról	balekkal/ balekkel

When there are no harmonic vowels in the root, front vowels are normally chosen for the suffixes:

(6)

		(adess.)	(delat.)	(instr./com.)	(allat.)
address	cím	címnél	címről	címmel	címhez
hand	kéz	kéznél	kézről	kézzel	kézhez
film	film	filmnél	filmről	filmmel	filmhez

But about fifty roots take the back variants of alternating suffixes:

(7)

		(as)	(pl.)	(ablat.)	(owner: 3 psg.)
bridge	híd	hídul	hidak	hídtól	hídja
arrow	nyíl	nyílul	nyilak	nyíltól	nyila
aim; target	cél	célul	célok	céltól	célja

The question as to how these neutral roots (henceforth called the híd words) should be described is another big problem within this field of investigation[4]. A third burning question is whether root harmony and suffix harmony ought to be described as a unitary process. In this article I will analyse and try to answer these problems. Furthermore the roundness harmony will be touched upon - the very limited assimilation process that lies behind e.g. one of the forms of the allative suffix: hez (cf. (2)).

3. THE MAIN PROBLEMS

3.1. The neutral exceptional roots

Roots without any harmonic vowels usually take front variants in the suffixes, as the harmony rule predicts. But as we saw, some neutral roots have instead back vowels in their suffixes.

Kiparsky (1968) divides neutralization into absolute neutralization - which is assumed to take place independently of context - and contextual neutralization, which shows up in a certain environment. He notes that only the existence of contextual neutralization has been proven.

A possible way of describing the place of the hid words in the harmony system is to accept absolute neutralization so that underlying back non-low unround vowels block suffix fronting (Kiparsky assumes that alternating suffixes have basic back vowels). Afterwards they merge everywhere with their front counter-parts.

Another solution is to assume that a non-phonological (diacritic) feature which is attached to the root conditions the harmony in suffixes.

A third alternative, which Kiparsky defends, is to introduce rule features to take care of the troublesome items and let phonological rules handle the majority of the words. Thus both cím and híd contain an /i:/ in the lexicon but híd is marked [- Vowel Harmony] so the suffix vowels do not change.

The preference for the first solution in analyses of numerous similar problems is said to depend on diachronic considerations only (for instance, there should have been - in Hungarian - unround back vowels that have merged with front vowels[5]).

If synchrony and diachrony could be described by the same rules, research in linguistics would no doubt have been easier. But children that acquire their native language "do not have the interests of linguistics at heart". Kiparsky further notes that "contextual neutralizations are reversible, stable and productive whereas the alleged absolute neutralizations are irreversible, unstable and unproductive".

Vago's (1973) main purpose is to show that the neutral vowels which govern back harmony are best described as underlying back vowels (ɨ→i, ɨ́→í, ʌ́→é). This view is referred to as the abstract solution. The exceptions are said to be systematic - all híd roots have just neutral vowels. Diacritics should not apply to systematic exceptions and accordingly not to the neutral exceptional roots.

Then Vago claims that he, Esztergar (1971) and Stong&Jensen (1971)- independently of each other - proved the rule feature analysis to be incorrect.

The evidence comes from the personal pronouns in other cases than the nominative and the accusative. Possessive suffixes are in these cases attached to the case markers, which thus act like stems. (The personal pronoun in its basic form is generally not used here, since it conveys emphasis and otherwise gives the same information as the suffix.) The following examples are taken from Vago:

(8)		(from)	(with)	(off)	(at)
I	én	tőlem	velem	rólam	nálam
you	te	tőled	veled	rólad	nálad
he, she, it	ő	tőle	vele	róla	nála
we	mi	tőlünk	velünk	rólunk	nálunk
you	ti	tőletek	veletek	rólatok	nálatok
they	ők	tőlük	velük	róluk	náluk

Vago concludes that the word-initial case forms display the underlying value for the feature back[6]. Some of the case morphemes moreover occur as verb prefixes, which agree in backness with the corresponding case stems. Another piece of evidence is the underlying frontness of the conditional -na/-ne - in the first person singular all verbs pick up the front variant.

So Kiparsky's assumption that all alternating suffix vowels are basically back must be rejected and with it his rule feature solution. (In the rule feature analysis híd+től would not change to hídtól - even if vowel harmony did not apply.)

Farkas (1979) presents some evidence against Vago's abstract solution. His informants are native speakers of Hungarian, living in Rumania. When they try to speak Rumanian, roots with harmony determining vowels that are similar to ı, í, Á do not behave like the híd words (but follow the general harmony pattern) and the vowels do not change as the absolute neutralization rule predicts.

Even more interesting is material from three speakers, showing that cél (aim; target) and derék (honest; brave; waist) may take front vowel variants of less common suffixes and back vowel variants of more common suffixes. Derivational suffixes almost always have the back vowel variants. E.g.: céltalan (aimless), derekas[7] (well), célnak, céltől, célba/célbe, dereka, deréktől. This is of course strong support for a nonabstract theory. Farkas notes that no roots with i or í show such a variation. I suppose the special behaviour of cél and derék depends on the alternation between á and é in suffixes. This relationship should strengthen the attraction towards true harmony as in the above examples.

We may now conclude this passage in the following way:

Absolute neutralization would constitute, as Kiparsky points out, a very small part of all neutralization processes. There is only disconfirming evidence - with the single exception of the deceptive simplicity of the solution. Farkas showed that there is a tendency - in words with a putative Á - towards true harmony in at least one Hungarian dialect.

The abstract solution is inferior for another reason: sounds that have not been heard in Hungarian for centuries are now treated as existing but latent, while the merger is repeated on and on.

Evidently a synchronic solution is preferable. The exact nature of the rules will be investigated in 4.2.

3.2. Neutral vowels

As we have seen, Vago assumes that e, é, i and í are the neutral vowels in Hungarian (the seven-vowel dialect). The criterion that Vago makes use of is that neutral vowels may appear in any morpheme together with front harmonic vowels - e.g.: sügér (perch) - as well as back vowels -e.g.: virág (flower) - without it sounding odd to a Hungarian.

Ringen (1978) vindicates for various reasons the claim that e is a harmonic vowel.

For instance, none of the híd words contain an e and the same goes for all invariable suffixes. Some of the mixed vocalic words with the neutral vowels i, í or é take only back suffixes - this is not true for those with e. She notes also that in certain native roots - like betyár (outlaw; scamp) - e co-occurs with back vowels. But all such co-occurrences are said to depend on the merger ë→e, so that the original form was bëtyár[8]. Ringen saves a solution with harmony as a unitary process moving from the first harmonic vowel in the root by saying that e is obligatorily exempted from harmony in roots.

Vago (1978) explains the reason e does not enter into any of the híd words (note, however, derék) is because those with Á are secondary developments and only found in two roots. As e alternates with a and there is no independent motivation for an abstract ʌ, e can never be invariable. The last objection is turned down because the list of focal exceptions would be too long.

Vago (1980) has even more criticism against Ringen's view. Against Ringen's assumption that no neutral, but in general all harmonic suffix vowels are sensitive to harmony, Vago counters with the fact that the suffix kor never alternates - though o is harmonic. But Vago himself (1978b) satisfactorily analyzes the suffix as an exception (low vowel lengthening does not appear before kor either), so the objection is not valid.

That é alternates with á in some suffixes is not a good argument either. The suffix vowels i and í are always invariant (e.g.: ni (infinitive ending) and ít (causative verbal ending)), while é sometimes is invariant (e.g.: é (a possessive derivational suffix, marking that the root is a possessor of something in the linguistic context)) but may also vary with á (e.g.: nál/ nél (adessive)). Examples (where the verb - unless otherwise stated - is given in its basic form, i.e.: 3 psg. in the indefinite conjugation of the present indicative): mosolyog (smile, i.e.: someone smiles) - mosolyogni (to smile); boldog (happy, delighted) - boldogít (make happy, favour); Kovács (family name) - Kovácsé (something that belongs to Kovács); a ház (the house) - a háznál (at the house). A reasonable interpretation of the facts is to regard á as the source in case there is an alternation, otherwise é.

But Vago has a reliable argument, namely that a number of roots govern back harmony but have an e in the last syllable (back vowels for the rest) - e.g.: maszek (self-employed), maszeknak (+dative/genitive), *maszeknek.

Let us summarize the results in the following manner:

Ringen's view that e is harmonic in Standard Hungarian has both advantages and disadvantages. If only neutral vowels may be skipped over in roots - as in (3) - then the fact that words like maszek condition back harmony is unexplainable. Vago, on the other hand, proposed that co-occurrence within native roots is characteristic of neutral vowels - here e acts as if it was neutral. But if e is regarded as neutral, it ought to be invariant in suffixes and then forms like *maszeknek would come up.Thus:

(9)	a) e is harmonic	b) e is neutral
	maszek+nek →	maszek+nek →
	*maszeknek	*maszeknek
	Root-e gives front vocalism to suffixes = no change.	Root-e is skipped over but neutral suffix vowels never change.

Facts actually seem to be a bit more complicated, even if Kiefer (1984) only mentions back suffix vowels (acc.and pl.) after this root. One native informant would spontaneously say maszeknak, but adds that the root is likely to vacillate in the speech of many Hungarians. maszek is furthermore considered a vacillating root in Papp (p. 167). Even so, one possible variant is underivable in both solutions - as shown in (9).

A neat way to solve the problem is to accept that e is ambivalent as regards neutrality. In roots it acts like a neutral vowel but in suffixes as a harmonic vowel. e is then typologically neither a neutral vowel nor a harmonic vowel.

In the dialects with ë the situation is much simpler. Sima (1980) says that [æ] does not co-occur with back vowels in native Finnish roots and the same goes for the eight-vowel dialects. Here the low front vowel is truly harmonic .

3.3. The domain of progressive harmony

Now we may investigate our next problem - which deals with the range of progressive harmony: whether the process affects all vowels or whether only suffix vowels harmonize as a result of it.

Kiparsky (1968) suggests that stem harmony should be described by morpheme structure conditions, because its exception classes differ from those of affix harmony. His premiss is right: there are disharmonic loanwords like parfüm (perfume) or zsonglőr (juggler) but suffixes on a line are never disharmonic in themselves, e.g.: (j)eitëkhëz/(j)aitokhoz (the harmony rule with the concomitant exceptions determines what alternant the root should take; the endings mark that the root is a possessed noun in plural in the allative case and whose owner is 2 ppl.). Vago (1973) accepts this use of MSC's, primarily due to the fact that stems normally do not alternate.

Kiparsky (1973a) refers to an input which arises from morpheme combination or by the application of a phonological rule as a derived input. It is shown that there are phonological rules in Finnish, Estonian, Swedish and Sanskrit which only apply to derived inputs. A new alternation condition is now formulated as:

(10) Neutralization processes apply only to derived forms.

Then ten cases for absolute neutralization - among them Vago (1973) are taken into account and the condition is accordingly altered to accept (preliminarily) such processes, iff they are automatic (applying to all occurrences):

(11) Non-automatic neutralization processes apply only to derived forms.

This is of interest here because Vago (1976) remarks an interesting difference between the forms of certain words: bokor (bush)/bokrok (k: pl.), tükör (mirror)/tükrök (pl.). In these words epenthesis and later harmony probably affect the singular and other forms. A similar difference is displayed in words like falu (village)/falvak (pl.) and tetű (louse)/ tetvek.

Vago concludes that harmony affects these singular forms because they are derived (in Kiparsky's sense). This is a good explanation, at least for the second group.

In an issue of Linguistic Inquiry, Jensen (1978), Phelps (1978) and Ringen (1978) criticize Vago's approach, while Vago (1978a) defends it.

Ringen's approach follows both from her assumption that e in roots is neutral - true, though not enough elaborated upon - and from rule (18), which I do not find explanatorily adequate.

Both Jensen and Phelps believe in abstract segments and place them even in roots like bika (bull) so that the harmony will be identical for roots and suffixes. To roots like Tibor - whose diminutive (with the derivative suffix -i) takes front suffixes (e.g.: Tibinél) - they add an exception feature.

But Vago points at iskola (school) - whose diminutive form isi acts like Tibi (it takes front suffixes). But isi is not used by all school children. It would be odd if both analyses were used in this case, so Vago rejects their solution. He also adds a stronger objection - some recent loan-words like sofőr (chauffeur) are disharmonic according to native judgement while words like bika are felt to be quite regular.

Jensen (1984) is an attempt within lexical phonology to show that harmony is also at work in roots. But his arguments are far from convincing.

The treatment of zero feature marking has a definite affinity to what I regard as a weakness in (18) - two groups of morphemes that should have been kept apart are combined.

Rules that ignore intervening neutral vowels are said to be unnatural, but such a rule makes - as we saw in the preceding section - a correct and important generalization (the want of which Jensen's solution does not compensate).

A vacillation was reported by Vago (1980) in the case of suffixes preceded by both a back-vocalic root and the suffix né (wife of). Jensen solves this by marking né (-), which is intended to mean optional variation between negative and neutral value for backness.

One native informant reports however that né is a truly neutral *ending* and that the morpheme-combination tanárné (wife of teacher) in careful speech takes back suffixes . Tanárné is actually rare, but the more usual papné (wife of priest) has only the possessive form papnéja (not papnéje) attested in ÉrtSz. The existence of occasional or possibly dialectal front suffixes might partly depend on the simultaneously semantic and phonetic similarity between né and nő (woman) - cf. tanárnő (female teacher). They are surely related, but né is not a free morpheme - which strongly suggests that it forms part of the suffixes. That the vowel is neutral allows on the other hand for both possibilities, since that property should prevent it from alternating (even if it were derived). The fact that its status as a suffix is comparatively vague presents itself as an explanation for e.g.: tanárnének. In the Standard dialects we should not expect to encounter such front forms to any extent. Hence it is possible to treat né as the other suffixes with neutral roots.

In short, the separation of root harmony and progressive harmony seems well motivated.

4. FORMALIZATIONS

4.1 Harmony conditions

Except for some newer and unassimilated borrowings, vowels from the two harmonic classes do not co-occur in any morpheme. Let us therefore assume that backness in vowels is normally a suprasegmental feature that covers every morpheme, so that the same value for backness is inserted in the matrices for all non-neutral vowels that are included.

Vowels in disharmonic words must on the other hand be specified for backness on the underlying level and this is of course more costly.

Before the harmony condition we need a condition that assigns the value [-back] to neutral vowels, namely:

(12)
$$\begin{bmatrix} \text{V} \\ \text{-round} \\ \text{-low} \end{bmatrix} \rightarrow [\text{-back}]$$

For the dialects which lack *ë* the neutral vowel condition has to be expressed like this:

(13)
$$\begin{bmatrix} \text{V} \\ \text{-round} \\ \left\{ \begin{matrix} \text{-low} \\ \text{-long} \end{matrix} \right\} \end{bmatrix} \rightarrow [\text{-back}]$$

Harmonic vowels may then be described by a condition[9] that the two dialects have in common:

(14)
$$\text{If:} \quad \begin{bmatrix} \text{V} \\ \text{0back} \end{bmatrix} \; / \; \begin{bmatrix} \text{+morpheme} \\ \alpha\text{back} \\ (\ldots \_ \ldots)_1 \end{bmatrix}$$
$$\downarrow$$
$$\text{Then:} \quad [\alpha\text{back}]$$

Notice that the opposite order (condition (14) before the condition that assigns frontness to neutral vowels) is possible but not desirable. The alternatives to the inputs in (12) and (13), respectively, demand more space and seem intricate in comparison. Evidently the former order is better off.

A mirror-image MSC which prohibits front round and back vowels from co-occurring would serve no purpose - we would have to insert separate values for every vowel in every morpheme and then the disharmonic ones must carry exception markings. One could naturally think of the first or the last vowel as trigger - but there is no evidence in either direction.

The vowels in *kaland* (adventure), *lát* (see), *orca* (cheek), *őz* (deer, roe), *görög* (Greek) and *rügy* (bud, burgeon) are thus underlyingly unspecified for backness, while the neutral vowels in *idő* (time; weather; tense (n.)) and *palacsinta* (pancake) are fronted by the earlier condition.

4.2 Harmony rules

These harmony rules are presented in Vago (1976):

(15)

$$\text{(m) VH: } [\text{+syll}] \rightarrow [\text{+back}] \;/\; \begin{bmatrix} \text{+syll} \\ \text{+back} \end{bmatrix} C_0 \left(\begin{bmatrix} \text{+syll} \\ \text{-back} \\ \text{-round} \end{bmatrix} C_0 \right)_1 \underline{\quad}$$

(16)

$$\text{(u) VH: } [\text{+syll}] \rightarrow [\alpha\text{back}] \;/\; \begin{bmatrix} \text{+syll} \\ \alpha\text{back} \end{bmatrix} C_0 \underline{\quad}$$

If (m) VH applies, (u) VH is blocked. This disjunctive order is just what might be expected if we accept Kiparsky's (1973b) Elsewhere Condition.

Ringen (1980) also presents two rules:

(17)

$$\begin{bmatrix} \text{V} \\ \left\{ \begin{matrix} \text{+low} \\ \text{+round} \end{matrix} \right\} \end{bmatrix} \rightarrow [\alpha\text{back}] / \begin{bmatrix} \text{V} \\ \alpha\text{back} \end{bmatrix} \left(C_0 \begin{bmatrix} \text{V} \\ \text{-low} \\ \text{-round} \end{bmatrix} \right)_0 C_0 \underline{\quad}$$

(18)

$$\begin{bmatrix} \text{V} \\ \left\{ \begin{matrix} \text{+low} \\ \text{+round} \end{matrix} \right\} \end{bmatrix} \rightarrow [-\alpha\text{back}] \;/\; \begin{bmatrix} \text{V} \\ \alpha\text{back} \\ \text{D} \end{bmatrix} \left(C_0 \begin{bmatrix} \text{V} \\ \text{-low} \\ \text{-round} \end{bmatrix} \right)_0 C_0 \underline{\quad}$$

Condition: obligatory when root contains only neutral vowels,
optional otherwise.

As starting-point we will use Ringen's first rule (slightly revised), which is quite general and corresponds well to the normal circumstances in the dialects with eight short vowels.

In section 3.1 the abstract solution was rejected. Ringen's diacritic feature theory is not satisfactory either. Collapsing e.g. *o* in *pozitív* and *í* in *híd* under the designation D seems counter-intuitive. It is like denying the real reason for vacillating roots.

If both methods have to be refused - what expedient may then be chosen?

Because of examples like those in (8), Kiparsky's rule feature theory had to be given up in subsequent theories. But that its failure is only illusory will be shown on this chart:

(19)				
Harmony	aludnál	nőnél		alszik: sleep
0Harmony	aludnék	nőnék	innék	nő: grow
-Harmony			innál	iszik: drink
-na/-ne: conditional		-l: 2 psg.		-k: 1 psg.

So the fifty exceptional stems are treated as inducing [-Harmony] (disharmony) - [+Harmony] is the unmarked case and not needed as specification in lexicon.

Chomsky and Halle (1968, p. 374) assumed that [- rule n] only meant non-application of rule n to a given so marked item. It is however clear that the motivation for their proposal is weak. Naturalness is not a conclusive argument, especially when it implies simplicity united with lack of counter-evidence.

We may assume that the extraordinary result - opposite to the expected - could be stored for the híd words in the lexicon as a negation on the rule. My hypothesis is that in a rule where a segment in the context has an α which determines in part the output, [- rule n] means a minus marking on the contextual α.

It is obvious that harmony fulfils the demands of the new Alternation Condition (11). Only the derived areas are affected.

Anderson (1980) notes that there are languages (Nez Perce, Luorawetlan, Diola Fogny and others) with two sets of vowels: dominant and recessive. Only if all underlying segments are recessive will the word contain recessive vowels. Otherwise all vowels will be dominant.

The typical harmony rule differs from (17) - but both imply that all vowels in the word will agree with regard to the relevant feature. Type (17) implies neither that harmony is non-automatic nor that it may only change derived forms. The limitation on harmony can be expressed by means of a marking on the changing segment:

(20) $$V^d \rightarrow [\alpha \text{back}] \;/\; \begin{bmatrix} V \\ \alpha\text{back} \end{bmatrix} \left(C_0 \begin{bmatrix} V \\ -\text{low} \\ -\text{round} \end{bmatrix}_N \right)_0 C_0 \begin{bmatrix} \underline{\quad\quad} \\ \begin{Bmatrix} +\text{round} \\ +\text{low} \end{Bmatrix} \end{bmatrix}$$

There is now no need for a marked harmony rule. A mechanism that counts the last neutral vowel in words like koncert (4) as determinant accounts for forms like koncertről. The mechanism is optional for roots like pozitív. We may use 'N⊆d' - where ⊆ is inclusion. It is to be read: ë, é, i and í in underived areas to the right of the last back vowel are not parts of N - instead they determine harmony.

Without this restriction a harmonic vowel is the determinant - cf. (3).

An example consisting of the same string with two derivations might be clarifying: poz**i**t**í**v+jΔ**i**+k+ról → pozitívjaikról (about their positives), pozitív+jΔ**i**+k+ról → pozitívjeikről (the symbol Δ designates a low short vowel which is unspecified for backness; functionally neutral vowels (i.e. parts of N) are printed in boldface). The root is unmarked in the first derivation - which works allright under the assumption that the scope of the rule (determinant with intervening and determined vowels) is maximal.

Earlier in this article we have seen that e in the seven-vowel dialects acts like a neutral vowel in roots but in suffixes as a harmonic vowel. The state of affairs can be incorporated into the rule[10]:

(21)

$$V^{d} \rightarrow [\alpha\text{back}] \,/\, \begin{bmatrix} V \\ \alpha\text{back} \end{bmatrix} \left(C_0 \begin{bmatrix} V \\ -\text{round} \\ \left\{ \begin{matrix} \text{d.}-\text{low} \\ -\text{back} \end{matrix} \right\} \end{bmatrix}_N \right)_1 C_0 \begin{bmatrix} \underline{\qquad} \\ \left\{ \begin{matrix} +\text{round} \\ +\text{low} \end{matrix} \right\} \end{bmatrix}$$

There is yet another problem. After the segment structure conditions have applied (vid. Stanley, 1967) á may change equally well into é and e, a may become ö or e, while e has two nearest back equivalents - a and á.

Here Vago (1974) makes use of adjustment rules, which Farkas (1979) gives in an informal manner.

Vago has a rule [a]→[ɒ] which accounts for the fact that when suffixes with underlying [æ] alternate, the vowel not only becomes back but round as well.

a-adjustment may be stated as:

(22)

$$\begin{bmatrix} V \\ +\text{low} \\ +\text{back} \\ -\text{long} \end{bmatrix} \rightarrow [+\text{round}]$$

An é-adjustment rule is needed for the á/é alternation: harmony applies - [a:]→[æ:] - and then [æ:] is raised to [e:].

Vago collapses e-adjustment (a lowering needed for e.g. roundness harmony in the seven-vowel dialects) and ê-adjustment into one rule, which will look like this:

(23)
$$\begin{bmatrix} V \\ -high \\ -back \\ -round \\ -\alpha long \end{bmatrix} \rightarrow [\alpha low]$$

(22) and (23) must however be replaced in this description because they are not synchronically motivated. In the next section alternative ways out will be described.

With the markings a on input and b on change vacuous rule application can be effectively hindered. This is achieved by a new condition:a≠b. Now the arrow always means change.

But this convention should rather be universal - a general property of every phonological rule, not necessary to mention on this low level. The alternative would be to accept vacuousity; the rules need anyhow no adjustment in themselves.

4.3 Roundness harmony and replacing the adjustment rules

To account for derivations like gyümölcs+ei+tök+höz→gyümölcseitëkhëz (to your (pl.) fruits) and kert+öm+höz → kertëmhëz (to my garden) - as opposed to gyümölcs-ö-tök+höz→gyümölcsötökhöz (to your (pl.) fruit) and kert+ünk+höz→kertünkhöz (to our garden) - the following rule will do[11]:

(24)
$$V^{d} \rightarrow [-round] \;/\; [-round]\, C_0 \begin{bmatrix} \underline{\quad} \\ -high \\ -long \\ -back \end{bmatrix}$$

This rule is simpler than Vago's (1974) - adopted in Ringen (1980) and which mentions the feature lowness - and the fact that it applies vacuously (ë→ë and/or e→e) does not matter, as we saw in the preceding section.

In the case of unrounding in the dialects without ë, there are actually three possible outputs: i, e and é. Of the nonlabial vowels, the three just mentioned are most close to ö - that is, they have the greatest number of distinctive features in common with ö.

The eventual output - e - is low. This may be expressed in the rule by adding the feature [+low] immediately to the right of the arrow. But as the determinant may be non-low as well as low this expedient does not seem appealing.

If we consider two other rules with similar problems, the description will overall be better. The rules are harmony and low vowel lengthening (LVL).

The existence of the latter is defended in Vago (1978b). LVL is responsible for alternations like kefe (brush), kefét (accusative), keféje (poss. 3 psg.), keféjét; alma (apple), almát, almája, almáját.

The rule looks like this[12]:

(25) $$\begin{bmatrix} +\text{syll} \\ +\text{low} \end{bmatrix} \rightarrow [+\text{long}] \; / \; \_ + [+\text{segment}]$$

Here a may change to á or ó and e may change to é or ő. A closer look at the precedence for a certain feature over another in the rules might be revealing[13]:

(26)

Harmony:	á→é	e→a	a→e	Roundness:	ö→e	ö→e
	long	long	low	(dialects	(not i)	(not é)
	low	round	round	without ë)	high	long
LVL:	a→á	e→é			low	low
	low	back				
	round	low				

For example: a and ö have another value for round than e, while a and e in contrast to ö share the same value for low. Because a becomes e in the harmonic feature exchange we put it this way: low matters more than round.

It is easy to see that there is a general pattern: round is below low while the other features are on the top.

We may further assume that the precedence for a feature over another is indicated in the feature order on the Hungarian vowel matrix. As an example the harmonic change a→e is given below[1]:

(27) a → [-back]

	a	á	e	é	ë	ö	ő	o	ó	u	ú	ü	ű	i	í
long	−	+	−	+	−	−	+	−	+	−	+	−	+	−	+
back	+		−		−	−		+		+		−		−	
high	−		−		−	−						+		+	
low	+		+		−	−									
round	+														

Now each segment which on a certain step agrees with the changing segment **s** may proceed, except when the output and **s** have different values. Then the reverse value is preferred so that any segment which exhibits it passes over to the next feature, the next link in the chain.

Instead of two diachronic rules we have an order between the features in the vowel matrix.

Another solution consists of introducing a principle, that predicts what will happen:

If the value for feature **f** changes through a rule, then let us assume that the participants in that rule (or a possible extension of it) move towards symmetric pair relations (where two elements pair solely with each other).

Consider the following description:

(28) Harmony | RH (seven-vowel dialects) | LVL (eight-vowel dialects) | LVL (seven-vowel dialects)

```
Harmony        RH (seven-vowel    LVL (eight-vowel    LVL (seven-vowel
                  dialects)           dialects)           dialects)

  o←→ö           ü←→i                ë←→é             ő←→ö↔é↔i←→í
  ↙↗              ↙↗                  ↙↗                 ↙↗
  a               ö↔e                 e                  e
  ↖↘              ↖↘                  ↖↘                 ↖↘
    e             ő←→é                  á                  á
  ↙↗                                  ↙↗                 ↙↗
 á                                    a                  a
  ↖↘                                   ↖↘                 ↖↘
    é                                 o←→ó               o←→ó
  ↙↗
 ó←→ő
```

A dotted line marks total similarity except for f; a continuous line marks that the two segments are separated by two features.

e→é in LVL (eight-vowel dialects) does not seem predictable at first. Harmony surely plays a role here - it might by the way be ordered after LVL (two endings are sensitive only to the former rule, vid. Vago (1978b)). We may also consider that á is back and so not neutral - a word like szőke (blond) would be disharmonic whenever a derived input followed it.

Everything else follows neatly from the principle - a→e (harmony), ö→e (RH) and a→á (LVL) because primary pair relations block the alternatives. e→a makes the sufficient addition for the creation of a secondary pair relation. e therefore repels á, which consequently becomes é.

e→é in the seven-vowel dialects does not improve our conviction about the principle's reliability since the change is already accounted for.

When it comes to deciding between the two devices, things are more uncertain. They are both rather simple but one may ask whether any one of them has an anchorage in reality. Yet they differ beneficiently from (22) and (23) by being based on the structure of present-day Hungarian speech only.

5. SUMMARY

Evidence has been given for a division of harmony - on the morpheme level it is determined by condition (14), while a rule makes the suffixes harmonize uniformly.

Neutral vowels are as stated above invariable in suffixes. Together with some other pieces in the puzzle this leads us to the conclusion that in the seven vowel dialects e belongs to neither the harmonic nor the neutral group, but acts in roots as a neutral vowel and in suffixes as a harmonic vowel. é, i and í are thus the proper neutral vowels, while front round and back vowels are harmonic. In the eight-vowel dialects, ë is neutral and e is harmonic everywhere.

All affected (changing) vowels are automatically defined as derived inputs. This makes the nature of harmony more transparent and the marker d obviates the need for a second rule.

It has been shown that a rule feature analysis is possible for the híd words, if Chomsky and Halle's convention to account for lexically determined rule non-application is given up and replaced by a new convention.

Instead of diachronic rules, two kinds of ways to derive correct surface forms have been suggested.

Though my aim with this article has been to give an alternative to the predominant historicistic view through functioning and unexpensive methods, it may well be that the main finding is how e in dialects without ë reacts to harmony.

Footnotes

1. It may be noted that the vowels here designated as a, ë and e are transcribed in IPA as [ɒ], [e] and [æ], respectively. In the seven-vowel dialects, e is more close - something like [ɛ] but with a special timbre.

2. Harmony causes the following alternations: a/e, á/é, o/ö, ó/ő, u/ü, ú/ű.

3. The tendency to choose the front variants is positively related to the number of final neutral vowels but in individual cases the harmony is not predictable. Kombiné (slip, undergown) thus takes back suffixes (or, more likely, vacillates), while oxigén (oxygen) takes the front variants. The two words differ from each other with respect to harmony in spite of the fact that their vowel sequences are identical.

4. Harmony occurs between all sorts of roots and endings, thus: zöld (green), zöldebb (greener), korrekt (correct), korrektebb/korrektabb (more correct), zsír (grease), zsíros (greasy), föld (earth), földes (earthy), fúj (blow), fújtam (I blow+preterite), nyit (open), nyitottam (I open+pret.).

5. But Kiparsky (1968) observes that Finnish vowel-initial suffixes are back after neutral monosyllabic roots. This rule may originate from the same source as the harmonic behaviour of the híd words.

6. Underlying forms of the case suffixes are: -nek (dative/genitive), -be (illative), -ben (inessive), -ből (elative), -ra (sublative), -ról (delative), -hoz (allative), -nál (adessive), -től (ablative), -vel (instrumental/comitative) and -ért (causal/final). All except the last one alternate. Some of the case stems are slightly different from the corresponding suffix forms, e.g.: belől- → -ből.

7. The form is dërëkas in the eight-vowel dialects - the basic form is dërék. Derék is the only bisyllabic stem in the exceptional híd group.

8. This already mentioned merger is by all means quite natural; its existence is confirmed in Sima (1980).

9. Adopting the Elsewhere Condition - vid. Kiparsky (1973b) - we may state the changing segment simply as [V].

10. In a lexical phonology framework, d might be replaced by the number of that level (or stratum) where alternating suffixes are adjoined.

11. The marker d is needed for the root eszköz (means; tool; instrument).

12. Low final vowels are never lengthened before a root morpheme, e.g.: körtefa (pear tree), faláb (wooden leg).

13. From now on until the beginning of (27) long, high etc. are to be taken as abbreviations for longness, highness etc.

References

Anderson, Steven R. (1980): "Problems and perspectives in the description of vowel harmony", in Robert M. Vago, ed. (1980).

Anderson, Steven R. and Paul Kiparsky, eds. (1973): A Festschrift for Morris Halle, Holt, Rinehart and Winston, New York.

Chomsky, Noam and Morris Halle (1968): The sound pattern of English, Harper and Row, New York.

CIFU 6 = Congressus Internationalis Fenno-Ugristarum: 6 (1980), Turku.

ÉrtSz. = A magyar nyelv értelmező szótára [Hungarian explanatory dictionary], I-VII, Budapest, 1959-1962.

Esztergar, Marianne (1971): A generative phonology of nouns and vowel harmony in Hungarian, Dissertation, University of California, San Diego.

Farkas, Donka (1979): "How abstract is Hungarian phonology ?", in Papers from the 15th Regional Meeting Chicago Linguistic Society.

Fujimura, Osamu, ed. (1973): Three dimensions of linguistic theory, TEC, Tokyo.

Jensen, John T. (1978): "Reply to 'Theoretical implications of Hungarian vowel harmony'", Linguistic Inquiry 9:1, 89-97.

Jensen, John T. (1984): "A lexical phonology treatment of Hungarian vowel harmony", Linguistic Analysis 14:2-3, 231-253.

Kiefer, Ferenc (1984): Magyar-Svéd szótár [Hungarian-Swedish dictionary], Akadémiai Kiadó, Budapest.

Kiparsky, Paul (1968): "How abstract is phonology ?", in Osamu Fujimura, ed. (1973).

Kiparsky, Paul (1973a): "Abstractness, opacity, and global rules", in Osamu Fujimura, ed. (1973).

Kiparsky, Paul (1973b): "'Elsewhere' in phonology", in Steven R. Anderson and Paul Kiparsky, eds. (1973).

Nyman, Alf (1960): Tankefel och tankefällor, Gleerup, Lund.

Papp, Ferenc (1975): A magyar főnév paradigmatikus rendszere [The paradigmatic system of the Hungarian noun], Akadémiai Kiadó, Budapest.

Phelps, Elaine (1978): "Exceptions and vowel harmony in Hungarian", Linguistic Inquiry 9:1, 98-105.

Ringen, Catherine O. (1978): "Another view of the theoretical implications of Hungarian vowel harmony", Linguistic Inquiry 9:1, 105-115.

Ringen, Catherine O. (1980): "A concrete analysis of Hungarian vowel harmony", in Robert M. Vago, ed. (1980).

Saussure, Ferdinand de (1960): Course in general linguistics, Peter Owen Limited, London.

Sima, Ferenc (1980): "Zur Characteristik der Ungarischen und der Finnischen Vokalharmonie", Pars VI, 258-266, in CIFU 6 (1980).

Stanley, Richard (1967): "Redundancy rules in phonology". Language 43, 393-436.

Stong, Margaret T. and John T. Jensen (1971): Two constraints in phonology, Syracuse University and Mc Gill University, unpublished.

Vago, Robert M. (1973): "Abstract vowel harmony systems in Uralic and Altaic languages", Language 49, 579-605.

Vago, Robert M. (1974): Hungarian generative phonology, unpublished Doctoral dissertation, Harvard University, Cambridge, Massachusetts.

Vago, Robert M. (1976): "Theoretical implications of Hungarian vowel harmony", Linguistic Inquiry 7:2, 243-263.

Vago, Robert M. (1978a): "Some controversial questions concerning the description of vowel harmony", Linguistic Inquiry 9:1, 116-125.

Vago, Robert M. (1978b): "The lengthening of final vowels in Hungarian", in Ural-Altaische Jahrbücher 50, 144-148.

Vago, Robert M. (1980): "A critique of suprasegmental theories of vowel harmony", in Robert M. Vago, ed. (1980).

Vago, Robert M., ed. (1980): Issues in vowel harmony, Benjamin, Amsterdam.

Vowel intrinsic pitch in Standard Chinese

Shi Bo and Zhang Jialu

Institute of acoustics, Academia Sinica

ABSTRACT

We investigated whether an intrinsic pitch (IP) effect occurs in Standard Chinese (SC) and if it exists how IP and pitch level interact with each other. The fundamental frequencies (F0) of each 9 Chinese vowels at different tonal points were measured in three positions: (1) in a monosyllable, (2) in the word-initial and (3) the word-final position of a disyllabic word. The test items (400 monosyllables and 509 disyllabic words) were embedded in a frame sentence and uttered by 5 male and 5 female informants. The results show that the characteristics of IP are to be found in all different tones of SC in spite of the fact that those tones have different tonal configurations. Further, the higher the relative pitch value, the larger the difference in F0 among the vowels. The IP differences are reduced in word-final position. These results suggest a new hypothesis.

INTRODUCTION

Intrinsic pitch (intrinsic F0) describes the influence of tongue height of vowels on the F0-value associated with them: high vowels have higher average F0-values than low vowels when other factors are kept constant.

A great deal of research has been devoted to the analysis and quantification of intrinsic pitch in several languages: English (Crandall,1925; Taylor, 1933; House and Fairbanks,1953;

Lehiste and Peterson, 1961), Italian (Ferrero et al., 1975), Danish (Petersen, 1978), Japanese (Nishinuma, 1977), French (Di Cristo, 1985), German (Neweklowsky, 1975), Greece (Samaras, 1972), Taiwanese Chinese (Zee, 1978), Yoruba (Hombert, 1977), Serbo-Croatian (Ivić and Lehiste, 1963), Itsekiri (Ladefoged, 1964), and Chinese (Connel et al., 1983). IP has also been observed when vowels were sung at the same pitch (Ewan, 1979. Personal communication by C.K.Chang).

Various experimental conditions were applied in these studies. In the early experiments, isolated "real" words (such as in Peterson and Barney, 1952) as well as "nonsense" words (such as in House and Fairbanks, 1953) were used. The segmental environments (i.e. consonantal context) were carefully controlled. Later, the test words were embedded in a frame sentence (as in Lehiste and Peterson, 1961: "Say the word ___ again."). The effects of prosodic environment on IP were also studied. Petersen (1976) reported that the magnitude of IP in stressed syllables is larger than in unstressed syllables. Similar results were obtained for Italian accent/nonaccent words (Ferrero and al., 1975). All of these studies generally showed similar results except Umeda's (1981) which reported that there were no consistent IP effects in a 20-min reading by two speakers. In order to investigate whether IP effects occur in connected speach, Ladd and Silverman (1984) compared test vowels (in German) in comparable segmental and prosodic environments under two different experimental conditions: (1) a typical laboratory task in which a carrier sentence served as a frame for test vowels; (2) a paragraph reading task in which test vowels occurred in a variety of prosodic environments. It was shown that the IP effect does occur in connected speech, but that the size of the IP differences is somewhat smaller than in carrier sentences. They pointed out that Umeda's finding was questionable because she apparently had not made any attempt to control for the prosodic environment of the vowels that were measured. In a recent study, Shadle (1985) investigated the interaction of IP and intonation in running speech. She examined the F0 of the vowels [i,a,u] in four sentence positions. The results showed a large main effect of IP that lessened in sentence final position. The study by Zee (1978) on Taiwanese Chinese showed that the IP also appears in a tone language, and that its magnitude is less for lower tones. In his study only two contrasting tones, high tone and low tone were analyzed.

However, none of these studies were concerned with the roles of pitch level and the position in the word in affecting intrinsic pitch. The main goal of the present experiment was to get a general idea about the effect of intrinsic pitch in Standard Chinese. The effect was to be studied as a function of the following variables: (1) pitch level (in different tones); (2) position in disyllabic words (word-initial and word-final).

METHOD

The material consists of two parts, 400 monosyllables and 509 disyllabic words. The monosyllables consist of all possible arrangements of consonants and vowels in Chinese, each arrangement having all four tones. Among them there are 279 "real" monosyllabic words and 121 "nonsense" words. The disyllabic words consist of one test syllable (a vowel preceded by an initial consonant) and one matched syllable. The matched syllable was chosen in such a way that the test vowels could be compared in a similar segmental environment and the same tonal surroundings. Examples are <u>fā</u>huà/<u>fū</u>huà; wēi<u>bā</u>/wēi<u>bō</u>/wēi<u>bī</u>; <u>tú</u>jǐng/<u>tí</u>xǐng, (The test syllables are underlined). Of the test syllables 273 were in word-initial and 236 in word-final position. As many combinations of two tones as possible were involved in this part.

In order to make all test items occur in the same phonetic environment and approach the situation of connected speech, all the monosyllables and disyllabic words were embedded in the frame sentence / Wǒ dú _ zì./ "I utter the character _." and / Wǒ dú _ _ zhè ge̊ cí./ "I utter the word _ _." respectively.

10 speakers (5 males and 5 females) of Standard Chinese were recorded. Before the recordings each speaker was given a short training period . A natural style was aimed at. The test materials were read once by each speaker in an acoustically treated room.

The records were fed into a Visi-Pitch (model 6087) for the extraction of F0. The counter provides a digital display of F0 for sustained vowels while the cursor allows the user to determine the F0 of any point on the pitch curve shown on the screen with +1 Hz accuracy.

Fig.1 shows the measuring points of F0. They are: for high tone (T1) the middle point T1; for rising tone (T2) the lowest point T2-1 and the highest point T2-2; for dipping tone (T3) the starting point T3-1 and the lowest point T3-2; for falling tone (T4) the highest point T4-1 and the lowest point T4-2.

As a first step in our analysis we only cared about average IP differences between vowels but ignored the differences between consonantal context and interspeaker variabilities. The statistical method was a one-way analysis of variance (with speakers and consonantal environments as a repeated measure).

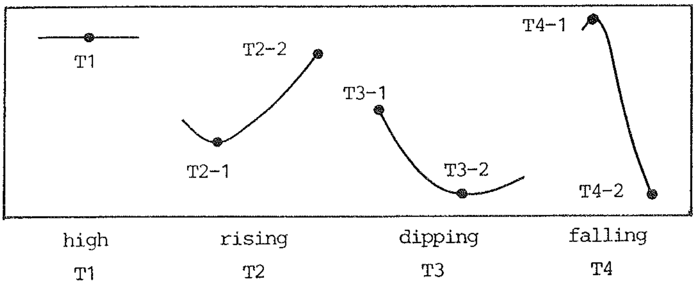


Fig.1 Measuring points of fundamental frequency

RESULTS

1. Vowel intrinsic pitch in four tones

The data which will be analyzed in this section were derived from 400 monosyllables. The intrinsic F0-values for each of 9 vowels at different tonal points are given in Table 1 in which the data are mean values averaged across consonants, for 5 males and 5 females respectively. The vowels were arranged in order of F0-value (based on the data in Table 1 averaged across 10 speakers) in Table 2. The relative mean F0 differences between low vowel [a] and the remaining vowels are shown in Table 3 (see `monosy.` parts).

The data mentioned above permit us to make the following observations: 1). At points T1, T2-2, and T4-1 the F0-values of high, middle and low vowels go from high to low and the F0 differences between high and low vowels are 10-30 Hz; 2). At points T2-1 and T3-2 a high vowel also has a higher F0 except that the F0-value of [o] is a bit higher than that of [ɿ] and [i]. This is also shown graphically in Fig.2 (see ·—·) The data at these five points show that Chinese also exhibits the influence of intrinsic pitch.

At points T3-1 and T4-2 the situation is more complex. In the measurements we found considerable inter and intra-speaker variability for F0-values at T3-1. For point T4-2, the main problem is that at the end of T4 the energy is very low and the

Table 1. Mean intrinsic F0-value for each of the 9 Chinese vowels at different tonal points, derived from 400 monosyllables, averaged across consonantal contexts, and for 5 males and 5 females respectively

sex	vowel (IPA)	number of occurrences (each tone)	F0-value (Hz)						
			T1	T2-1	T2-2	T3-1	T3-2	T4-1	T4-2
male	i	12	175	118	167	113	89	197	97
	ɿ	3	181	122	171	116	90	208	99
	ʅ	4	179	116	169	115	90	195	101
	y	6	180	119	175	115	90	197	101
	u	19	181	117	168	112	90	206	105
	e	19	164	114	156	114	88	187	101
	o	18	168	117	160	116	90	184	100
	ɚ	1	170	116	170	122	88	178	100
	a	18	154	111	151	108	83	175	97
female	i	12	291	205	265	219	169	312	180
	ɿ	3	302	206	271	214	172	326	182
	ʅ	4	295	200	264	216	168	319	192
	y	6	300	209	278	219	171	318	176
	u	19	307	209	289	218	172	335	184
	e	19	289	202	270	215	170	315	183
	o	18	278	200	270	213	170	310	183
	ɚ	1	302	200	274	209	161	314	182
	a	18	276	198	255	227	171	302	187

Table 2. Vowels ordered from high F0 to low F0 (derived from 400 monosyllables, averaged across 5 male and 5 female speakers and consonantal contexts)

F0	T1	T2-1	T2-2	T3-1	T3-2	T4-1	T4-2
high	u	u	u	y	u	u	ʅ
↓	ɿ	y	ʅ	ɿ	ɿ	a	u
	y	ɿ	y	u	y	i	e
	ʅ	ʅ	ɿ	i	o	ʅ	a
	i	i	i	o	ʅ	ɿ	o
	e	o	e	ʅ	i	u	ɿ
	o	e	o	e	e	o	y
low	a	a	a	a	a	e	i

Table 3. Relative F0 differences (Δ F0) between the vowel [a] and the remaining 7 vowels at different tonal points in three conditions: monosyllable, word-initial and word-final, (data were averaged across consonantal contexts and all speakers)

point	position	i	ɿ	ʅ	y	u	e	o	mean 2**
		F0 difference (Hz)							
T4-1	monosy.	16	28	18	19	32	12	8	19
	word-i.	16	5	11	15	22	19	18	15
	word-f.	8	10	9	6	16	13	2	9
	mean 1*	13	14	12	13	23	16	9	14
T1	monosy.	18	27	22	25	30	12	8	20
	word-i.	20	24	30	23	21	21	20	23
	word-f.	11	23	8	19	12	15	13	14
	mean 1	17	25	20	23	21	16	14	19
T2-2	monosy.	11	16	11	21	23	8	10	14
	word-i.	22	21	11	21	17	9	7	15
	word-f.	13	-	2	15	14	4	3	10
	mean 1	15	18	8	19	18	7	7	13
T2-1	monosy.	7	9	4	9	8	4	4	6
	word-i.	12	21	12	7	10	3	2	9
	word-f.	11	-	8	5	11	0	2	6
	mean 1	10	15	8	7	10	2	2	7
T3-2	monosy.	2	4	2	4	4	2	3	3
	word-i.	5	7	6	6	5	3	-2	4
	word-f.	2	5	3	1	3	-1	1	2
	mean 1	3	5	4	4	4	1	1	3

* mean for 3 positions
** mean for all 7 vowels

A negative value indicates that the F0-value of vowel was lower than that of vowel [a]

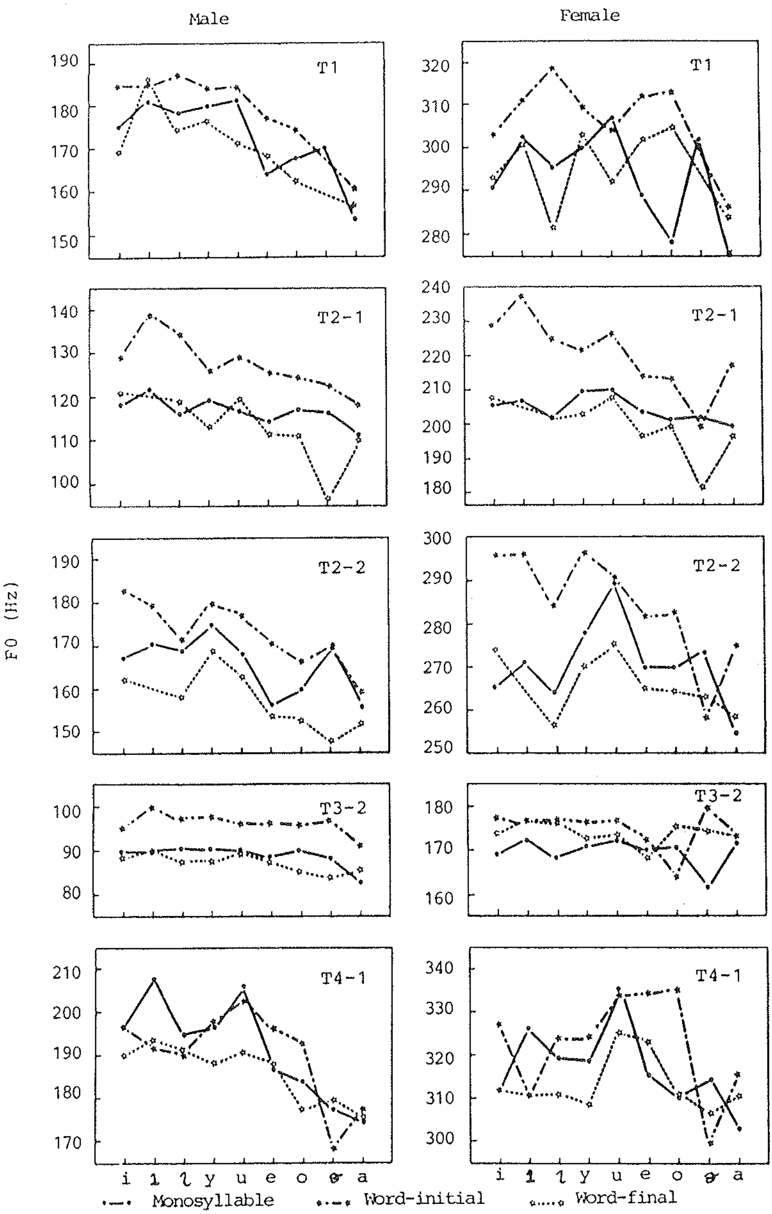


Fig.2 Mean F0 for the vowels, plotted as a function of tongue height, averaged across consonants and for 5 males and 5 females respectively

periodicity is not good enough to permit precision in measurements. As a result there is no consistent influence of IP at these two points.

2. Effect of word-position on IP

When the test syllables were in the disyllabic words there are additional factors influencing F0. There is some modification in F0 caused by the adjacent tone. The data in Fig.2 (★ and ☆) and table 3 (see `word-i.` and `word-f.` parts) show that, though meaning and tonal environment are not separated in the data, the effect of intrinsic pitch still occurs regardless of whether the test vowels were in word-initial or word-final position. However the magnitude of IP was reduced in word-final position. This reduction appears to be related to a lowering of F0 in this position (in Fig.2, the curves derived from the word-final position are the lowest ones in most cases).

3. Interaction of intrinsic pitch with pitch level

Fig.2 shows F0-values of 9 simple vowels as a function of the tongue height associated with them. In each part of Fig.2, from left to right, the tongue height of the vowel goes from high to low and it is accompanied by a drop in F0, which reflects the effect of IP. But the curves at different tonal points have different slopes. It indicates that although each vowel appears to be associated with an intrinsic F0 at a certain tonal point, the differences of intrinsic F0 across the vowels vary from point to point. This variance can be seen more clearly in Table 3, in which from top to bottom, the mean relative F0 differences (ΔF0) between the low vowel [a] and the remaining vowels become smaller and smaller as F0 drops. The ΔF0 at points T1 and T4-1 (high F0) are much larger than those of point T3-2 (lower F0).

Going a step further, there is little difference in ΔF0 between the males and the females in spite of the fact that the F0 of the females is higher than that of the males. It indicates that the magnitude of ΔF0 is directly proportional to some kind of relative pitch value rather than to the absolute F0-value. In tone languages, "tonal value" and "tonal register" are often used to discribe the relative relationships of pitch values. If we call the absolute F0-minimum as F0(min) and F0-maximum as F0(max), then the tonal value T(p) (in Oct.) for F0(p) (in Hz) is the binary logarithm of the quotient of F0(p) and F0(min). When F0(p) is equal to F0(max), the T(max) is the tonal register.

Thus

Tonal value: $T(p)=\log_2(F0(p)/F0(min))$ Oct.

Tonal register: $T(max)=\log_2(F0(max)/F0(min))$ Oct.

The ΔF0 between i-a and between u-a are plotted as a function of the normalized tonal value (T(p) divided by T(max)) in

Fig.3. The "●" represents the averages over i-a and u-a, across the males and the females. It is obvious that the higher the tonal value, the larger the ΔF0. In other words, the IP is more marked in the high frequency region of the tonal register than in the lower one. But the slopes of the two curves of the females turn negative when the normalized tonal value is bigger than 0.8. It seems that when the F0-value goes beyond certain limits, the direct proportional relation between ΔF0 and F0 will no longer be tenable. This suggests that it might be worth while to study intrinsic pitch in a larger F0 dynamic range such as in singing.

Table 4. Comparison of FD* for different languages

FD of		FD of Chinese male	FD of Chinese female
Italian~	23.3	25.2 (T1)	23.5 (T1)
		24.5 (T4-1)	22.6 (T2-2)
Danish	16.5	14.5 (T2-2)	19.6 (T4-1)
Greece	11.9		
English	10.5		
French	9.8		
German	9.6		
Italian**	9.2	7.2 (T2-1)	9.1 (T2-1)
Japanese	9.0	7.2 (T3-2)	

* FD=FH-FL FH: mean F0-value of high vowels [i,u,y]
FL: mean Fo-value of low vowels [ɛ,œ,ɔ,a]

~ accent

** unaccent

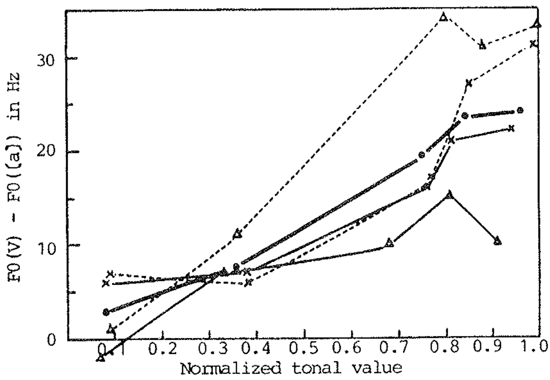
4. Comparison

For the purpose of comparison, the term "FD" was defined as the difference between the mean F0-value of high vowels [i,u,y] (FH) and the mean F0-value of low vowels [ε,œ,ɔ,a]) (FL), i.e. FD=FH-FL. The FD-values of some other languages (quoted from Di Cristo, were compared to the FD-values of SC in Table 4. It seems that the magnitude of IP of SC at different tonal points just varies between those of other languages. Italian and Chinese are similar with respect to the fact that for Italian the FD-value of an accented syllable is larger than the FD-value of an unaccented one, and for Chinese the FD-value for a higher tone is larger than the FD-value for a lower tone.

The mean ratios of [i] and [u] to [a] as established in the present study and in a number of other investigations are given in Fig.4. It is clear that the data of the present study are in good agreement with the results of other studies.

DISCUSSION

There have been various hypotheses for the cause of IP. Taylor's "dynamogenetic irradiation hypothesis" (1933) appears to be the earliest attempt. He thought that "the most plausible explanation of the vowel pitch triangle appears to be that it results from dynamogenetic irradiation of the tongue tension to the vocal cord musculature." But it was obviously



Males: X Females: ▲ (i): —— (u): ---

Averages over males and females, (i) and (u): ●—●

Fig.3 Mean F0 differences between (i,u) and (a) are directly proportional to the normalized tonal value

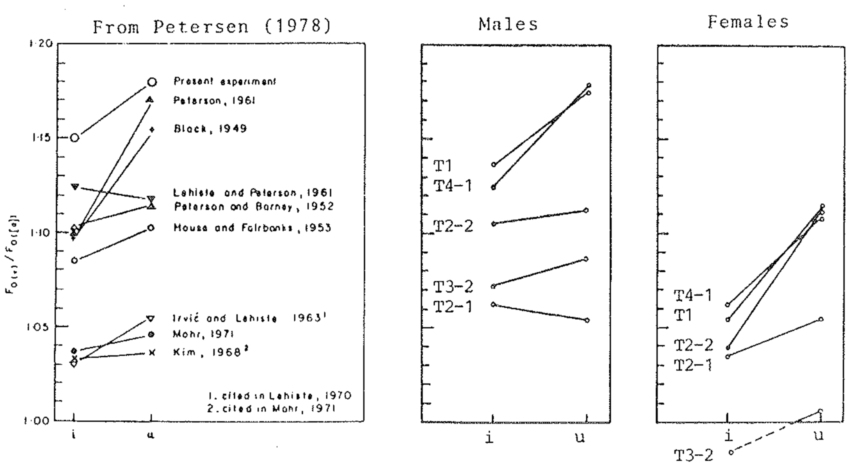


Fig.4 Mean ratios of (i) to (a) and (u) to (a), based on the data from the present and a number of other investigations

questionable.

There is a source/tract coupling hypothesis (See i.e. Flanagan and Landgraf, 1968; Atkinson, 1973) based on the assumption that the IP could be caused by acoustic interaction between the first vowel formant and the vibration of the vocal folds. Since higher vowels have lower first formants than low vowels, the acoustic interaction should be greatest for high vowels whose first formant frequencies are closer in frequency to F0. However, the results of the experiments made by Beil (1962) and Ewan (1979) contradict this hypothesis. And the relationship between [i] and [u] does not tally with what the hypothesis predicts either. We usually find that the first formant F1 of [i] is lower than that of [u]. According to the hypothesis [i] should be expected to have a higher F0 than [u]. Nevertheless, the present and a number of other investigations show the opposite relation, i.e. a higher F0 in [u] than [i] (See Fig.4).

The [i]-[u] relation we just discussed also contradicts another hypothesis -- Mohr's "pressure hypothesis" (1971). Mohr assumes a pressure of air to be built up behind the supraglottal constriction, which reduces the airflow through the glottis and, consequently, the rate of vocal fold vibration. The greater the distance between the constriction and the glottis, the longer it takes for the air pressure to be established behind the constriction and hence for the F0 to drop. If this explanation were true, a higher F0 should be expected in [i] than in [u]. Thus, the results of the present study support neither source/tract coupling nor pressure hypotheses.

Of the various hypotheses, it seems that the tongue pull theory has received the greatest attention. The early tongue pull hypothesis (Ladedoged, 1964) supposed that the tongue, when raised to produce high vowels, pulls the hyoid bone and the larynx upwards, thus resulting in an increased vocal-fold tension which in turn leads to a higher F0. But this explanation is contradicted by the fact that the hyoid/larynx position always seems to be lower in [u] than in [a]. Ohala (1973) modified the tongue pull hypothesis. He thought that the increased tongue pull in high vowels gives rise to increased vertical tension in the vocal folds through the mucous membrane and other soft tissues without involving the hyoid bone and the hard tissues of the larynx. In support of this explanation, it appears that there is a positive correlation between ventricle size, which is assumed to reflect vertical tension in the vocal folds and tongue height and intrinsic F0 of vowels. The tongue pull hypothesis has been expanded further by Ewan (1975). In addition to tongue pull--or rather, lack of tongue pull-- Ewan proposes a "tongue retraction/pharyngeal constriction component" to account for the low intrinsic F0 of low vowels. Ewan suggests that the low F0 of low vowels, which are also assumed to involve a tongue retraction or pharyngeal constriction component, is caused by the soft tissues being pressed downwards in the direction of the larynx and thus increasing the vibrating mass of the vocal folds, which results in a decrease in F0.

But few of these hypotheses attempt an explanation of the "nonlinearity" in IP. In Chinese the higher the tonal value, the larger the IP difference; in Italian, the accented syllables display greater IP than unaccented ones (Ferrero et al., 1975); deaf speakers often exhibit a larger than normal IP which may be related to a higher than normal average F0 (Bush, 1981). IP is reduced in final sentence position with a lowered F0 (Shadle, 1985). The common point is that a larger IP difference seems always correlated to a higher F0. Moreover, the variation of tonal characteristics due to syntactic and semantic factors is much larger at the tonal roof than at the tonal floor (Bannert, 1984). So a larger variation of F0 always corresponds to a higher F0. And this sort of nonlinearity is relative to a <u>within</u>-subject variation (i.e. it does not mean the female should be expected to have a larger IP difference than the male because of a higher voice). There was a simpler explanation that general relaxation (as in an unaccented phrase-final position) may reduce intrinsic F0. But it is contradicted by the evidence against vowel neutralization in that "relaxed" sentence position (Shadle, 1985).

Here, we try to give a probable interpretation from the point of inherent nonlinearity of the vocalis muscle itself. According to Ohala's theory (1973) the tongue pull gives rise to increased vertical tension in the vocal folds through the mucous membrane and other soft tissues. We could assume that there must be a series of deformations in the mucous membrane and the soft tissues, and finally in the vocalis muscle itself thus causing increased tension. The relationship between the tension T and the elongation x of the vocalis muscle can be approximately expressed as (Fujisaki et al., 1981):

$$T = a\ \exp(bx)$$

The incremental tension per unit elongation, as given by $\gamma T/\gamma x$, is obviously greater at larger values of x which generally correspond to higher F0-values. In other words, the same incremental elongation due to the tongue pull could cause a larger increase in tension T, thus leading to a larger F0 variance at high F0 than at low F0. However, it must be emphasized that this is only a probable conjecture. The reliable evidence for the interpretation should be based on physiological data. Last, we think that if this kind of nonlinearity in the production of speech could be confirmed, it would be helpful for a better understanding of the similar nonlinearity found in the perception of speech.

ACKNOWLEDGMENTS

We would like to express our thanks to Prof. Eva Gårding and Prof. Maa Dah-You for their valuable suggestions and critical reading of the early versions. We are also greatly indebted to many colleagues for assistance, for encouragement, and for guiding us to relevant literature: David House, Li Meng, Li

Zi-yin, Lü Shi-nan, Qi Shi-qian, Jan-Olof Svantesson, and Paul Touati.

REFERENCES

Appelman, D. R. 1967. The Science of vocal pedagogy (Indiana U. P., Bloomington).

Atkinson, J. E. 1973. Aspects of intonation in speech: Implications from an experimental study of fundamental frequency. PhD dissertation, Univ. of Connecticut, Storrs, CT (unpublished).

Bannert, R. 1984. Towards a model for German prosody. Working Papers 27: 1-36. Dept. of Linguistics and Phonetics, Lund Univ..

Beil, R.G. 1962. Frequency analysis of vowels produced in a helium-rich atmosphere. J. Acoust. Soc. Am. 34: 347-349.

Bush, M. 1981. Vowel articulation and laryngeal control in the speech of the deaf. PhD thesis, MIT (unpublished).

Connel, B. A. et al. 1983. Experimental evidence of interaction between tone and intonation in Mandarin Chinese. J. Phonetics 11: 337-351.

Crandall, J. B. 1925. The sounds of speech, Bell Syst. Tech. J. 4: 586-626.

Di Cristo, A. 1985. De la microprosodie a l'intonosyntaxe. Publications Université de Provence.

Ewan, W.G. 1975. Explaining the intrinsic pitch of vowels. Paper presented at the Fifth California Linguistics Association Conference, San Jose, May 4, 1975, 1-9.

Ewan, W. G. 1979. Can intrinsic vowel FO be explained by source/tract coupling? J. Acoust. Soc. Am. 66: 358-326.

Ferrero, F. E. et al. 1975. Some acoustic and perceptual characteristics of the Italian vowels. Paper presented at the VIIIth Int. Congr. Phon. Sci. (Leeds), mimeographed.

Flanagan, J. and Landgraf, L. 1968. Self-oscillating source for vocal-tract synthesizers. IEEE Transactions on Audio and Electroacoustics AU-16, 57-64.

Fujisaki, H et al. 1981. Analysis of pitch control in singing. In K.N. Stevens and M.Hirano (eds.), Vocal fold physiology. Univ. of Tokyo press, 347-363.

Hombert. J-M. 1977. Consonant types, vowel height and tone in Yoruba. Stud. Afri. Linguistics 8: 173-190.

House, A. and Fairbanks, G. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. J. Acoust. Soc. Am. 35: 84-92.

Ivic, P. and Lehiste, I. 1963. Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom književnom jeziku. Zbornik za filologijy i lingvistiku 6: 33-73.

Ladd, D. R. and Silverman, K.E.A. 1984. Vowel intrinsic pitch in connected speech. Phonetica 41: 31-40.

Ladefoged, P. 1964. A phonetic study of West African languages. In West African Language Monograph, Series I (Cambridge U. P., Cambridge).

Lehiste, I. and Peterson, G. E. 1961. Some basic considerations in the analysis of intonation. J. Acoust. Soc. Am. 33: 419-425.

Lubker, J. et al. 1977. Vowel fundamental frequency and tongue height. J. Acoust. soc. Am. 62: S16-17(A).

Mohr, B. 1971. Intrinsic variations in the speech signal. Phonetica, 23: 65-93.

Neweklowsky, G. 1975. Spezifische dauer und spezifische tonhöhe der vokale. Phonetica, 32, 1: 38-60.

Nishinuma, Y. 1977. Contribution à l'étude prosodique du Japonais. Thèse de 3e Cycle. Univ. de Provence.

Ohala, J. 1973. Explanations for the intrinsic pitch of vowels. Monthly Internal Memorandum, Phonology Laboratory, Univ. of California, Berkeley (January), 9-26.

Petersen, N. R. 1976. Intrinsic fundamental frequency of Danish vowels. J. of Phonetics (1978) 6: 177-189.

Peterson, G. E., and Barney, H. L. 1952. Control methods used in a study of the vowels. J. Acoust. Soc. Am. 24: 118-127.

Samaras, M. 1972. Influence de l'entourage consonantique sur les variations de la fréquence laryngienne des voyelles du grec moderne, Bull. de l'Inst. de Phon. de Grenoble, I: 57-66.

Shadle, C. H. 1985. Intrinsic fundamental frequency of vowels in sentence context. J. Acoust. Soc. Am. 78: 1562-1567.

Taylor, H. C. 1933. The fundamental pitch of English vowels. J. Exp. Psychol., 16: 565-582.

Umeda, N. 1981. Influence of segmental factors on fundamental frequency in fluent speech. J. Acoust. Soc. Am. 70: 350-355.

Zee, E. 1978. The interaction of tone and vowel quality. UCLA work. Pap. Phonet. 41: 53-67.

Understanding measure and comparative sentences through Prolog

Bengt Sigurd

Measure sentences are e.g."Sven is 12 years old",Swedish: "Sven är 12 år gammal","The boat is 12 feet long", Swedish: "Båten är 12 fot lång", "The stone has a weight of 12 kilos", Swedish:"Stenen har en vikt av 12 kilo", "How much is the car?", Swedish:"Hur mycket kostar bilen?". Such sentences are part of the core grammar and play an important role in the use of natural language in communication. They have attracted the interest of linguists because they indicate how languages express fundamental cognitive dimensions and many linguists have noted that "old" in "How old is Sven?" does not presuppose that Sven is old, while "How young is Sven?" presupposes that Sven is young. It has also been noted that the sentence "Sven is old " means that Sven is old for a man and that even the absolute use of adjectives implies a comparison with a standard. Comparative constructions such as "Sven is five years older than Lisa " have added to the interest in measure adjectives. There is a large body of literature on adjectives and comparatives(cf.Rusiecki,1985). These works will not be reviewed in this paper, which aims at illustrating how such facts appear in a computer program which can "understand" measure and comparative sentences and answer corresponding questions.

This paper will present the basic syntactic and semantic features of Swedish measure and comparative sentences. The Swedish expressions will be translated and comparison with the equivalent English expressions will thus be possible. The computer program, written in Prolog, is constructed with this background. Various limitations have been introduced in order to get a program which can be used for demonstration purposes.

The program can parse a wide range of measure and comparative sentences. If it is asked to accept a sentence it adds the fact to its data base and returns "Jaha" (OK). If it is asked to answer a question it returns an answer if one is available.The program knows the equivalence of sentences such as "The boat is 30 feet long","The boat has a length of 30 feet","The length of the boat is 30 feet" and "The boat measures 30 feet" and is able to use accepted information in answers to several types of questions. It understands that it is reasonable to call a person who is over 60 old. The program understands that Sven is older than Bill if it has learnt that Sven is older than Lisa and Lisa is older than Bill. It understands that Lisa is 5 years younger than Sven if it has learnt that Sven is 5 years older than Lisa. It understands that Sven is double as old as Lisa if it has learnt that Lisa is half as old as Sven. In some variants of the program explanations of the conclusions are also added.

The program may be of interest to those who construct expert systems, where variants of measure sentences often play an important role. The fragment discussed can perhaps be embedded in a more comprehensive system. Some theoretical and psycholinguistic issues will be touched upon as the program forces the designer to decide where different types of linguistic and encyclopedic knowledge is to be put. The choice of logical representations is also an interesting theoretical issue. Several have to be chosen in order to handle the problems and pecularities discussed above. One might suggest that e.g. only the unmarked representation is stored in the memory ("Sven is bigger than Lisa" not "Lisa is smaller than Sven") and that inferences always, or preferably, are done on unmarked (positive) representations. This idea can easily be implemented in a version of the program to get a model which can explain why subjects can verbalize the unmarked adjective faster than the marked one(big before small) as has been found (Schriefers,1985).

The differences between Swedish and English evoke some comments and raise questions about measure and comparative sentences in the languages of the world - typological questions which will not be treated in this paper.

The syntax of measure sentences

Four main types of measure sentences can be distinguished: The adjectival type (Sven är fem år gammal),the nominal type (Sven har en ålder av fem år),the genitive type (Svens ålder är fem år) and the verbal type (Sven väger trettio kilo;lacking for the dimension age).

The syntactic structure of the first type can be illustrated by the following representation:

np(Sven),cop(är),pf(mf(num(fem),unit(år)), a(gammal)), where pf means predicative phrase, mf means measure phrase, a means (measure) adjective, and the other abbreviations should be evident . The characteristic feature is the copula and the adjective and in the Prolog program only the present tense "är" (is) and some common adjectives are used. The sample subject nouns are chosen not to create any variation in agreement.If e.g."barnet"(the child) is also included the ending -t would have to be added to the predicative adjective as "barnet " has neuter gender.("Barnet är tre år gammalt").It is quite possible to handle such problems in the Prolog definite clause grammar formalism used, but for our purpose it is important not to clutter up the program by the additional machinery.

One equivalent question is:

pf(mf(hur), a(gammal)), cop(är),np(sven)?(How old is Sven?), and an answer could be:"(Sven är) fem år (gammal)." It would also be possible to answer such a question without a measure phrase by saying:"Han är mycket gammal."(He is very old.) This case is not covered by our program. Another equivalent question is:

pf(mf(num(hur många),unit(år)), a(gammal)),cop(är),np(sven), but the only acceptable answer in that case is a numerical answer substituting for "hur_många"(how many), e.g."(Sven är) fem(år gammal)". The usual way of putting this quesions is in fact only:"Hur många år är Sven?" where "gammal" is deleted, probably as being redundant when "år"(years) is mentioned.

Sentences without measure phrases are similar to this type. But as many linguists have noted the meaning of some measure adjectives is quite different if used with a measure phrase. "Sven is five years old" does not mean that he is old, and that is why it is not too strange to continue:"He is not old". For basic dimensions there is an adjective which can be used in this neutral way;an exception is the dimension weight. As in English one cannot say:"Sven är fem kilo tung, nor "fem kilo lätt."(Sven is five kilos heavy/light). One has to say that:"Sven väger mycket"(Sven weighs a lot) or "Svens vikt är hög"(Sven's weight is great). "Sven är fem kilo tung" is strange, but it has an interpretation, which we might want our program to be able to handle.(A version does).

The second type of measure sentences is the nominal type, which may be illustrated by:

np(Sven),vp(har),np(en,mn(ålder),av,mf(fem,år))

where the characteristic feature is "har" plus the measure noun (mn) followed by a measure phrase.

The equivalent question is "Vilken ålder har Sven?" (What age is Sven?). This construction is slightly formal.

The third type of measure sentence is characterized by a subject where the bearer of the property is in the genitive and the property is denoted by a measure noun. It can be illustrated by:

np(Svens,mn(ålder)), cop(är),mf(fem,år). The equivalent question is "Vilken är Svens ålder?" It is not possible to say "Vilken ålder är Svens?" which would rather be an equivalent to the strange English sentence "Which age belongs to Sven?". It is, however, also possible to say (more informally) "Vad är Svens ålder?" A native is hardly aware of which kind of measure construction he uses.

There is a related nominal construction:"Sven har hög ålder"(Sven has high age) and a related genitive construction:"Svens ålder är hög"(Sven's age is high). We note that hög(high),låg(low) can be used with some dimensions e.g.ålder,kvalitet,kvantitet,höjd,hastighet(speed).

The fourth type of measure construction is characterized by the fact that the verb denotes the dimension, but there are only a few such verbs available. The construction can be illustrated by:

np(Bilen),vp(kostar),mf(8000,kronor)

The verb "kostar"(cost) denotes the price dimension, "väger" denotes the weight dimension (not "tynger", which means press down)."Rymmer","innehåller"(contains) denote the volume dimension, but other measure verbs are hard to come by. Some that come to mind e.g."mäter"(measure),"räknar"(count) seem often to have less acceptability and they are not associated with only one dimension. It seems possible to answer the question "Hur stor är salen"(How big is the hall?) by "Den mäter 1000 kvadratmeter", but also by e.g. "Den räknar 1000 stolar". It is also possible to take "stor" to indicate volume and answer "Den rymmer 500 personer".(It can take 500 persons, It can seat 500 persons). The verb "omfattar" would fit in a sentence such as "Den omfattar 1o sektioner"(It includes 10 sections). It is possible to answer the question "Hur lång är båten?" (How long is the boat?" by "Den är 30 fot lång", but also by "Den mäter 30 fot". The question "Hur bred är båten?" (How wide is the boat?" can possibly also be answered by a phrase containing the verb "mäter" e.g. "Den mäter två meter" (It measures two meters). The English verb measure seems to carry similar stylistic values.

There is a related construction without a measure phrase:"Bilen kostar mycket" (The car costs much), where the word "mycket" is an equivalent of "a lot". It is sometimes possible to use the verb alone (in a pregnant sense):"Det kostar", meaning "That costs a lot". In the same way some dimensional nouns can be used in a pregnant sense, e.g."Sven har tyngd" (Sven has weight), but in that case the meaning is abstract:Sven is of importance/carries weight.

The price dimension can also be denoted by "betingar". It may also combine with the noun:"Bilen betingar ett högt pris /5000 kronor"(The car costs a lot/5000 kronor).

Comparative sentences

Superlative sentences cannot normally take measure phrases as illustrated by the following sentences and ungrammatical expansions: "Sven är äldst av pojkarna"(Sven is (the) oldest of the boys), *"Sven är 5 år äldst av pojkarna"(Sven is 5 years (the) oldest of the boys). It is, however, possible to use a non-numerical modifier as e.g in "Sven är klart äldst av pojkarna"(Sven is clearly/by far the oldest of the boys). Superlative sentences are not included in the fragment of Swedish covered by our program, but it is interesting to speculate about the reasons why natural language(s) - at least not Swedish and English - do not offer superlatives for numerical quantification of the difference between an object and the group of objects it is singled out from on the basis of a certain dimension. If Sven is 5 years older than a certain group of boys this fact is rendered by a comparative construction where he is not considered to be a member of a group:"Sven är 5 år äldre än pojkarna"(Sven is 5 years older than the boys).

There are different types of comparative construction. Some types will be discussed below, but all are not covered by the Prolog program. We may distinguish between adjectival types, where the comparative forms discussed by so many linguists appear, e.g. "Sven är 5 kilo tyngre än lisa" (Sven is 5 kilos heavier than Lisa), verbal types, e.g. "Sven väger 5 kilo mer än Lisa" (Sven weighs 5 kilos more than Lisa), and nominal types e.g. "Svens vikt är 5 kilo högre än Lisas" (Sven's weight is 5 kilos greater than Lisa's), "Sven har en fem kilo högre vikt än Lisa" (Sven has a 5 kilos greater weight than Lisa).

Comparative sentences may also be divided into equality (comparative) sentences e.g. "Sven är lika tung som Per" (Sven is equally heavy as Per, Sven is as heavy as Per) and difference (comparative) sentences. Difference sentences may be divided into additional (and subtractional) sentences, e.g. "Sven är 5 kilo tyngre än Lisa" (Sven is 5 kilos heavier than Lisa) and multiplicational (divisional) sentences, e.g. "Sven är två gånger så gammal som Tim, Sven är dubbelt så gammal som Tim,Sven är två gånger (dubbelt) äldre än Tim" (Sven is two times (twice) as old as Tim). The divisional comparative type is illustrated by "Tim är hälften så gammal som Sven" (Tim is half as old as Sven). It is also possible to combine these types and verbalise an euqation as in "Sven är dubbelt så gammalsom Rut, minus 2 år" (Sven is double as old as Rut, minus 2 years), but such complex constructions will not be handled by our demonstration program.

The syntactic structure of an evaluative comparative construction without a measure phrase quantifying the difference can be illustrated by the following sentence.

```
np(Sven),cop(är),   compar(äldre),pp( prep(än),np(Lisa))
   Sven      is              older           than     Lisa
```

Equivalent questions are e.g. "Är Sven äldre än Lisa?" (Is Sven older than Lisa?), "Vem är äldre än Lisa?" (Who is older than Lisa?).

The syntactic structure of a related construction where a measure phrase is included can be illustrated by the following sentence.

```
np(Sven),cop(är),mf(num(5),unit(år)), compar(äldre)
   Sven      is          5     years            older
   pp(prep(än) np(Lisa))
          than     Lisa
```

There are several ways of asking equivalent questions depending on how many words are moved to the front of the sentence:"Hur många år är Sven äldre än Lisa?"(How many years is Sven older than Lisa?), "Hur många år äldre är Sven än Lisa?" (How many years older is Sven than Lisa?), "Hur många år äldre än Lisa är Sven?" (How many years older than Lisa is Sven?). English seems to have the same stylistic options. The unit(år) has to go with the question word (hur många), which supports the measure phrase (mf) as a constituent. The preposition(än) has to go with the noun(Lisa), which supports the prespoitional phrase as a constituent. But the comparative(äldre) can be moved with the measure phrase or be stranded with the prepositional phrase and this fact speaks against assuming a constituent including the comparative form and the prepositional phrase.

Multiplicational comparative constructions can be given a syntactic representation as illustrated by the following sentence.

```
np(Sven),cop(är),gf(num(5),gånger),adj(äldre)
   Sven       is          5  times        older
   pp(prep(än) np(Tim))
          than    Tim
```

With dubbelt (double) the construction illustrated by "Sven är dubbelt så gammal som Tim" is generally used, but in other cases both the construction with the comparative form and the construction "så adj som" are grammatical. This paper does not attempt to pinpoint all such idiosyncratic facts, however.

Equivalent questions support the constituent analysis suggested above. One may ask:"Hur många gånger är Sven äldre än Tim?"(How many times is Sven older than Tim?), "Hur många gånger äldre är Sven än Tim?"(How many times older is Sven than Tim?), "Hur många gånger äldre än Tim är Sven?"(How many times older than Tim is Sven?). An answer could be:"Dubbelt så gammal" (Double as old), "Tre gånger så gammal" (Three times as old) or "Tre gånger äldre" (Three times older). The wording of the answer is not dependent on the wording of the question, only on the meaning of the question.

Synonymy relations

As noted above there are often (almost) equivalent adjectival, nominal and verbal expressions. The following table lists the most important cases, where there are at least two word classes to choose from when a dimension is to be denoted together with a measure phrase.

Verb	Adjective	Noun
-	gammal(old)	ålder(age)
mäter(measures)	lång(long)	längd(length)
rymmer(contains)	stor(big)	volym(volume)
räknar(counts)	stor(big)	storlek(size)
väger(weighs)	-	vikt(weight)
-	hög(high)	höjd(height)
-	djup(deep)	djup(depth)
-	bred(wide)	bredd(depth)
-	tjock(thick)	tjocklek(thickness)
-	varm(warm)	temperatur
kostar(costs)	-	kostnad,pris(cost)
betingar(costs)	-	kostnad,pris(cost)

There are very few verbs which have a dimensional meaning and can be combined with a measure phrase. And some are ambiguous as noted. There exist only a few adjectives which can occur with a measure phrase and have a dimensional meaning in ordinary language. Dimensional nouns can always be formed, especially in more specific or scientific language. We might say that an object has flexibility 5, the colour resistence 3 etc.This is done in consumers tests. But even if scalar dimensions are thus devised one can only use the noun, not the equivalent adjective. One may say that something has a flexibility of 5 flex units, but hardly that it is 5 (flex) units flexible or that it flexes 5 units.

Logical representations

The logical representation of measure sentences with measure phrases is l(X,D,N,U) or l(X,D,E) , where X,D,E,N,U are variables. X is the object having the property. D is e.g.age, the name of the property(dimension). E indicates an evaluation and takes on the value more(+),indicating more than the standard for such objects or less(-), indicating less than the standard for such objects. The variable N covers numbers and the variable U covers units of different type. e.g. year, meter. A typical sentence with a measure phrase e.g. Sven is 5 years old would be represented by l(sven,age ,5,years). Time is disregarded.
The sentence Sven is old would be represented by the formula l(sven,age,more),where the evaluation variable is set at more meaning that Sven has a high age (but the number and unit variables are unknown). Different formats are thus used for the two types of sentences, and normally evaluation is incompatible with a specification of number and units. It is sometimes possible,although a bit strange,to give both an evaluation and a number of units and say e.g. Sven is fifty years young, indicating that Sven is young for his age and such cases can be handled if two facts are added to the data base, which can be done in various ways. But if this way of construction is permitted and considered as grammatical the grammar will be overgenerating.

The following table indicates how some common sentences will be represented in our logical format.

Sentences	Logical form
Sven är 2 meter lång.	l(sven,height,2,meter)
Sven är 50 år gammal.	l(sven,age,50,år)
Sven väger 70 kilo.	l(sven,weight,70,kilo)
Bilen kostar 5000 kronor.	l(bilen,cost,5000,kronor)
Lisa har en ålder av 40 år.	l(lisa,age,40,år)
Bilen är hög.	l(bilen,height,more)
Bilen är lätt.	l(bilen,weight,less)
Per är 70 år ung.	l(per,age,70,år),l(per,age,less) (in some versions of program)
Pers ålder är 70 år.	l(per,age,70,år)

Certain questions ask for numerical details, e.g."Hur många år(gammal) är Sven?" and they have to answered by calling the longer logical form and giving the values of the last two variables. Even the question "Hur gammal är Sven?" has to be answered by giving the values of the last two variables. The typical cases when the evaluation variable is to be looked for are yes/no questions such as "Är Sven gammal?". An answer could be "Ja" and if numerical values are available the answer could be expanded into "Ja, han är 75 år".(Not done in this program).Other questions of this type are:"Är bilen dyr?" (Is the car expensive?), "Kostar bilen mycket?" (Is the car expensive?).

Logical representations of comparative sentences are established similarly. We will represent "Sven är äldre än Lisa" by l(sven,age,more,lisa) or more generally l(X,D,E,Y), where X and Y are the two objects compared (X is in focus), D is the dimension of comparison, E is the evaluative variable, which may be more(with age representing older), or less (with age representing younger). This representation could,in fact, also be used for absolute cases, such as "Sven är gammal"(Sven is old), if we assign a standard to Y.

For the sentence "X är lika gammal som Y" (X is as old as Y) the representation l(X,age,more,like,Y) is used and for the sentence "X är lika ung som Y" (X is as young as Y) the representation l(X,age,less,like,Y) is used. The difference between those two sentences is that with "ung"(young) the persons are pressupposed to be young, while with "gammal" (old) no evaluation is implied (gammal being the neutral unmarked term). Maybe this presupposition should rather be represented as a separate fact, however.

The quantified comparative sentences need more complex representations. For the additional case illustrated by "Sven är 5 år äldre än Lisa" (Sven is 5 years older than Lisa) we write l(sven,age,more,5,years,lisa), where the forth and fifth arguments are the numbers and the units, respectively. Changing more into less allows us to represent the antonyms (older:younger,heavier:lighter,etc.). The case where the difference is given in multiplicational terms is illustrated by the following case:
l(sven,age,more,5,times,lisa), which is the logical translation of "Sven är 5 gånger så gammal som/äldre än Tim" (Sven is 5 times as old as/older than Tim).

The logical representations used are ad hoc and it is not clear how the should be integrated in a full semantics.

Some implications and inferences

Speakers of English draw the conclusion that Sven weighs 50 kilos, that his weight is 50 kilos and that he has a weight of 50 kilos if they have learnt that he weighs 50 kilos. They do not, in fact, consider these conclusions as consclusions rather as synonymy relations. This is a reason for treating them in the way it is done in our program, where all these expressions are represented by (translated into) the same logical representation. Once one of the expressions has been mentioned the others are automatically true and questions based on the other formulations are answered in the affirmative. The choice between adjectival, nominal or verbal expression does not involve any logical inferences in our program. The borderline between verbal synonymy and logical inferences is of theoretical interest, but it will not be discussed in this paper.

A genuine inference may be illustrated by deriving "Lisa is younger than Sven" from "Sven is older than Lisa". We may call this inference the "antonym comparative inference". In our Prolog program such inferences are handled by variants of the following rules.

```
l(X,D,more,Y) :- l(Y,D,less,X)
l(X,D,less,Y) :- l(Y,D,more,X)
```

Similarly the inference Lisa is 5 years younger than Sven can be drawn from the fact that Sven is 5 years older than Lisa, if the following rules are included in the program.

```
l(X,D,more,N,U,Y) :- l(Y,D,less,N,U,X)
l(X,D,less,N,U,Y) :- l(Y,D,more,N,U,X)
```

These implications and inferences are handled in a special section of the program which can be extended further. Presently this section also includes an equivalence inference

```
l(X,D,more,like,Y) :- l(Y,D,more,like,X)
```

This rule states that Sven is as tall as Tim if Tim is as tall as Sven. If we want to cover that Sven is as small as Tim if it is known that Tim is as big as Sven we could add a rule with less instead of more. A special variable ranging over more,less could also be used.

Multiplicational inferences can be handled by the following rule.

l(X,D,more,1/N,times,Y) :- l(Y,D,more,N,times,X).

This inference is illustrated by:Tim is 1/5 as old as Sven if Sven is 5 times as old as Tim. It is to be noted that only the numbers (N) are inverted in such cases, not the antonyms. It is not correct to conclude that Tim is 1/5 as young as Sven as that would imply that Sven was known to be young.

An inference based on the transitivity of the dimensions involved may be expressed in the following way. l(X,D,more,Z) :- l(X,D,more,Y),l(Y,D,more,Z)

This simply means that if X has the dimension in a higher degree than Y and Y has it in a higher degree than Z than X (also) has it in a higher degree(more) than Z. The case where numerical values are included is a bit more complicated. The following rules covers e.g. the following case:If X is 5 years older than Y and Y is 5 years older than Z than X is 10 years older than Z. l(X,D,more,N,U,Z) :- l(X,D,more,M,U,Y),l(Y,D,more,P,U,Z),N=M+P

We are now entering algebraic operations and further elaborations could imply solving equations, which would definitely take us inte the world of Artificial Intelligence. The age of Sven may for instance be computed if it is known that he is twice as old as Per, who is 6 years older than Bo, who is half as old as Bill, who is 60. But we do not want to prepare our program for such computations. They can hardly be included in the set of "natural implications".

Certain inferences dealing with evaluative sentences may be established although they are "fuzzy". If a person is over 60 he may be called old, and if he is under 20 he may be called young, but this varies with cultures and times. Similarly a price may be called high and the object expensive if it is above a certain figure, but this figure also varies with cultures, times - and, of course, with the type of object.

Comments on the program

The program is designed to be short and simple.It is meant to demonstrate one way of treating the scalar and comparative constructions commented on by so many linguists. The program runs on a VAX/VMS Version V4.2 and uses a Prolog version which the Department of Linguistics Lund has obtained by courtesy of the department of Cognitive Sciences Sussex (Sussex POPLOG Prolog (Version 10),1985). I am indebted to Robin Cooper for teaching me basic Prolog and to Mats Andersson for improving my program.

The first part of the program handles the parsing of declarative and interrogative sentences of the types discussed above. The predicate sent takes four arguments: a variable which is d for declarative and q for questions, a syntactic tree structure, the sentence(list) to be parsed and the empty list. The rules show which types of declarative and interrogative clauses the program can handle. The rules are written in the definite clause grammar formalism(DCG).The syntactic representation is simplified and many of the usual categorial labels have been left out in this restricted grammar.

The lexicon rules allow the insertion of certain definite nouns, which are non-neuter not to complicate inflexion by agreement (neuter and plural). The genitive form of nouns is constructed by adding an -s ('s) and that is handled by some rules below nps. Certain dimensional nouns are allowed, and some dimensional verbs. Some numerals and all integers are allowed as num in the measure phrases and a number of scalar units illustrate what can be accepted by the grammar.

The meaning of the words are given by the predicate b (for betydelse,meaning). The meaning is given by English words. We distinguish the dimensional meaning of e.g. gammal, which is age ,b(gammal,age), from the evaluative meaning of gammal,which is b(gammal,age,more). Note that absolute and comparative forms are assigned the same meaning here.

The syntactic structure derived through the parsing is translated into logical representations,e.g. for comparative l(X,D,E,N,U,Y) where X is the object which has the property, D is the name of the property,E is the evaluation variable, N is the numeral value and U is the unit.Y is the compared object. The values of D and E are derived by the meaning relations (b).

The program cannot accept strange sentences such as "Sven är 10 meter gammal"(Sven is 10 meters old). The knowledge that certain units are associated with certain dimensions is considered to be encyclopedic and not to be included in the grammar rules. The natural place for these constraints seem to be in the translation into logical form, thus prohibiting surrealistic sentences as the one above.

Various implications are included in a special section.The program handles the fact that a person who is over 60 years old may be called old. The predicate exp inserts the value more as the E if the dimension is age and the number of years exceeds 60 (an approximation).

The interactive predicates allow adding of information to the data base. The predicate accept(X) parses the sentence, derives the equivalent logical clause and stores it. It returns the word "Jaha"(OK) as a signal. The lexical and syntactic details of the sentences are thus not stored, which is possibly a feature of psycholinguistic reality. The facts are stored without checking inconsisistences. The inference rules are used only in answering questions.

There are many different ways of asking, which can be seen in the different definitions of answer(X). A typical question is "Hur gammal är Sven?" and this is handled if answer([hur,gammal,aer,sven,?]) is typed to the computer. The program will then parse the sentence according to the syntactic rules and derive the value q, and the tree structure s(sven ,[aer],mf([hur]), gammal). The logical form l(sven,age,N,U) will be called,and at last the values of N and U will be printed. In some cases a longer answer is generated on the basis of the logical representation. If there is no information available, if nothing has been said which allows any conclusions, the answer will be "vet inte". If there is a contradictory fact the answer will be "nej".

The interactive module of the program includes a function TEXT(X) which prints a text (series of sentences) telling about X. It will give all possible, even synonymous, verbalisations (with adjective,noun,verb) of the facts stored in the data base. The text often gets a bit verbose, but the function TEXT allows checking the program and what is stored currently in the data base. The command LOGIC(X) prints the logical (semantic) representation of the sentence X. Questions do not have logical representations in the present system (See Engdahl, 1986, for suggestions).

The command ANSWER(X) parses the sentence and gives an answer.The command SAY generates a sentence according to the grammar, prints it and adds the corresponding fact to the data base. This command simulates the situation where a speaker says something and a listener stores it. The command ACCEPT(X) is somewhat more complex as the storing is followed by a backchannel item, which is "Jaha"(OK) if the fact is new and "Jag vet" (I know) if the fact is known already. The command ASK generates a question which is also printed. These commands can be combined in different ways in order to simulate conversation. One way of doing this is shown by the command DIAL. This command calles the other commands in order to get new declarative sentences, questions, backchannel items and answers. These commands can be used in more advacnced simulations of communication, where two parts really exchange information.

Conclusion. A 7-modular model of the language user

The programming language Prolog offers a fairly convenient means for writing programs which parse and "understand" sentences. It is also very suitable for inverting the process and generating sentences on the basis of the same rules.It forces the linguist to make his ideas about syntax, lexicon, morphology, semantics, pragmatics and logic more precise. The working of the program is a proof that the analysis is correct, but not that the solution is the only possible, that it is the best, that it reflects the psycholinguistic processes of a speaker/hearer or that it would be suitable for the rest of the language. As can be seen from our program even such a small fraction as the measure and comparative sentences requires a big program.

The construction of this program handling measure and comparative sentences has lead to the establishment of seven integrated modules. The whole model may be called Integrative grammar, because the grammar is integrated among the 7 modules needed in a full model of a language user.
1.Categorial combination rules. These rules tell which categories or individual words can combine. They have to be expanded with agreement conditions in order to handle concord phenomena, which the rules in the program do not. These rules and the handling of transforms of sentences, gaps etc have been the focus of much linguistic discussion during the last decades. The categorial rules of combination also build trees in which the structure may be made more regular ("deep structures") and they give information about sentences mode (declarative and question) in a special variable. These rules can be parsed by an builtin mechanism in Prolog, which is not discussed at all here. This module is traditionally called Syntax.

2.Word category rules.
These rules tell which words fit into the categories mentioned in the categorial combination rules. This module is traditionally called lexicon and is mainly a list in the program.

In addition to the lists there are rules which may be called category derivation rules. These rules state that if a certain word belongs to a certain category,a certain variant of it belongs to a certain other category. Thus the addition of an -s makes a noun which belongs to the category genitive, and the addition of an -are to an adjective makes a word which belongs to the category comparative. These general rules do not cover all cases and a number of words have therefore to be listed separately, e.g. the irregular comparatives tyngre, mindre etc.). The contents of this module has traditionally been treated under the title Word formation or (Derivational) morphology. They get a natural place in integrative grammar.

3. Word meaning rules. These rules state that a certain word (form) has a certain meaning, may decomposed as in the dimenional adjectives under discussion. Our program gives statements such as for gammal b(gammal,age,more), where b is short for betydelse/bedeutung. A traditional entry in a dictionary would give the information about an entry (a form) at one place. These pieces of information are separated in our program.

Just as word category rules are supplemented with word category derivational rules, the word meaning rules are supplemented with meaning derivational rules. The meaning of a comparative form is thus derived from the meaning of the corresponding adjectival form by special rules.

4. Natural language-to-Logic rules (NLL rules) One rationale for translating the syntactic and lexical information into logic representation is that inferring can be expressed much more economically, if it is done with logical representations instead of natural language (surface) sentences. We will not discuss all the problems of semantic representations here.

5. Inference rules

Inference rules can be divided into encyclopedic rules and (purely) logical rules. The encyclopedic rules state that a person who is above 60 may be called old, that a town (generally) has streets, that a car (generally) has wheels etc. Logical rules are illustrated by the transitivity rules in the program, rules telling that X has more of something than Z, if X has more than Y, and Y has more than Z, or that X has more than Z if X has more than Y and Y the same as Z.

It seems the inference rules can be elaborated much if the converse terms are taken into account fully. The relations between parent and child, sell and buy are often mentioned, but there are many other predicates of this type in languages and they can also be used in inference rules (See further Sigurd,1976), e.g. teacher-student, doctor-patient, lawyer-client, like-please, know-familar.

6. Data base (memory) All linguistic interaction presupposes that the speaker has a data base (memory) in which he stores what he learns and which is being consulted when he answers questions. In the present system facts are stored by means of the built-in predicate assert. How the memory is structured and accessed is a well-known psycholinguistic problem. Some reaction time experiments (Schriefers,1985) indicate that unmarked adjectives can be produced faster than marked (negative) ones, and this interesting facts can easily be reflected in the design of the data base. Another problem is whether negative facts also should be stored, as indicated by ideas in situation semantics (Cooper,1984).

7. Interactive (pragmatic) rules The interactive commands of the last section of the program simulates interactive actions, which are triggered by the mode signals in speech. A declarative sentence triggers a back-channel item (OK etc.) and a question triggers an answer. The structure of this section is elaborated in Sigurd(1986) and has been the object of much research under the heading speech act theory. The present system has a place for these linguistic features as well.

The conclusion to be drawn from the present Prolog program for a fragment of natural language is that all the 7 modules and their interaction have to be elaborated and studied taking into account theoretical,typological and psycholinguistics facts. Sample interaction with the program is illustrated in an appendix. Extracts of the program is also given in an appendix.(The whole program is available on request.)

References

Bierwisch,M.1976. Some semantic universals of German adjectives.Foundations of Language 3

Bolinger,D.1972. Degree words. The Hague:Mouton

Cooper,R.1984. Sentence negation in situation semantics. Proc. 20th annual meeting of the Chicago linguistic society

Engdahl,E.1985. The syntax and semantics of questions with special reference to Swedish.Dordrecht:Reidel

Gnutzman,C.1974. Zur Graduierbarkeit von Adjektiven im Englischen. Linguistische Berichte 31

Lafrenz,P.G.1983. Zu den semantischen Struktur den Dimensionsadjektive in der Deutschen Gegenwartssprache. Acta Universitatis Gothoburgensis. Göteborg

Quirk,R, S,Greenbaum, G.Leech and J.Svartvik.1985. A Comprehensive Grammar of the English Language.London:Longman

Rusiecki,J.1985. Adjectives and comparison in English. A semantic study. London:Longman

Schriefers,H.J.1985.On semantic markedness in language production and verification. Diss.Nijmegen

Sigurd,B.1967. Adjektivkomparationens morfologi. Arkiv för Nordisk Filologi 82:1-4 (1967)

Sigurd,B.1976. Converse terms in Swedish.In: N.E.Enkvist and V.Kohonen (eds),Meddelanden från stiftelsen för Abo Akademi forskningsinstitut nr 8.

Sigurd,B.1986. Computer simulation of dialogue and communication.In: F.Karlson (ed), Proc. of the 3rd conf. of Scand. comptational linguistics. Helsinki:Dept of Linguistics.

APPENDIX I

SAMPLE DIALOG (SIMPLIFIED, IDEALIZED)

U(ser): SVEN HAR EN ÅLDER AV 70 ÅR. (SVEN HAS AN AGE OF 70 YEARS)
C(omputer): JAHA (OK)
U: ÄR SVEN GAMMAL? (IS SVEN OLD?)
C: JA. (YES)
U: HUR GAMMAL ÄR SVEN? (HOW OLD IS SVEN?)
C: 70 ÅR. (70 YEARS)
U: VAD ÄR SVENS ÅLDER? (WHAT IS SVEN'S AGE?)
C: 70 ÅR. (70 YEARS)
U: SVEN ÄR ÄLDRE ÄN LISA. (SVEN IS OLDER THAN LISA)
C: JAHA. (OK)
U: LISA ÄR ÄLDRE ÄN PER. (LISA IS OLDER THAN PER)
C: JAHA. (OK)
U: LISA ÄR LIKA GAMMAL SOM HANS. (LISA IS AS OLD AS HANS)
C: JAHA. (OK)
U: ÄR SVEN ÄLDRE ÄN PER? (IS SVEN OLDER THAN PER?)
C: JA. (YES)
U: ÄR SVEN ÄLDRE ÄN HANS? (IS SVEN OLDER THAN HANS?)
C: JA. (YES)
U: ÄR PER YNGRE ÄN HANS? (IS SVEN YONGER THAN HANS?)
C: JA. (YES)
U: ÄR HANS LIKA GAMMAL SOM PER? (IS HANS AS OLD AS PER?)
C: NEJ. (NO)
U: HUR MYCKET VÄGER HANS? (HOW MUCH WEIGHS HANS?)
C: VET INTE. (DON'T KNOW)
U: LISA ÄR 35 ÅR GAMMAL. (LISA IS 35 YEARS OLD)
C: JAHA. (OK)
U: HUR MÅNGA ÅR ÄR SVEN ÄLDRE ÄN LISA? (HOW MANY YEARS IS SVEN OLDER THAN LISA?)
C: 35 år. (35 YEARS)
U: HUR MÅNGA GÅNGER ÄLDRE ÄR LISA ÄN SVEN? (HOW MANY TIMES IS LISA OLDER THAN SVEN?)
C: 1/2 GÅNGER. (1/2 TIMES)
U: PER ÄR HÄLFTEN SÅ GAMMAL SOM HANS. (PER IS HALF AS OLD AS HANS)
C: JAHA. (OK)
U: HUR GAMMAL ÄR PER? (HOW OLD IS PER?)
C: 17,5 ÅR. (17.5 YEARS)
U: ÄR PER GAMMAL? (IS PER OLD?)
C: NEJ. (NO)

U: HANS VÄGER MER ÄN LISA. (HANS WEIGHS MORE THAN LISA)

C: JAHA. (OK)

U: VILKEN VIKT HAR LISA? (WHAT WEIGHT HAS LISA?)

C: VET INTE. (DON'T KNOW)

U: ÄR LISA TYNGRE ÄN HANS? (IS LISA HEAVIER THAN HANS?)

C: NEJ. (NO)

U: TEXT(SVEN).

C: SVEN ÄR 70 ÅR GAMMAL. SVEN HAR EN ÅLDER AV 70 ÅR. SVENS ÅLDER ÄR 70 ÅR. SVEN ÄR ÄLDRE ÄN LISA. SVEN HAR EN HÖGRE ÅLDER ÄN LISA. SVENS ÅLDER ÄR HÖGRE ÄN LISAS.

U: SAY.

C: SVEN ÄR STARK. (SVEN IS STRONG)

U: MSAY.

C: SVEN ÄR FEM ÅR GAMMAL. (SVEN IS FIVE YEARS OLD)

U: CSAY.

C: SVEN ÄR FEM ÅR ÄLDRE ÄN PER. (SVEN IS FIVE YEARS OLDER THAN PER)

U: CONV.

C: SVEN ÄR STARK.
JAHA.
ÄR SVEN ÄLDRE ÄN LISA?
VET INTE.

/* SYNTAX */ APPENDIX II:Fragment of program

```
sent(d, s(X, [aer], Y)) --> np(X), [aer], a(Y), [.
].
sent(d, s(X, Y, i(Z))) --> np(X), mv(Y), i(Z), [.]
.
sent(d, s(X, [har], id(Y), Z)) --> np(X), [har], i
d(Y), mn(Z), [.].
sent(d, s(X, [aer], Y, Z)) --> np(X), [aer], mf(Y)
, a(Z), [.].
sent(d, s(X, [har], Y, Z)) --> np(X), [har], [en],
 mn(Y), [av], mf(Z), [.].
sent(d, s(X, Y, [aer], Z)) --> nps(X), mn(Y), [aer
], mf(Z), [.].
sent(d, s(X, Y, Z)) --> np(X), mv(Y),mf(Z), [.].
sent(d, s(X, [aer], Z, [aen], W)) --> np(X), [aer]
, compar(Z), [aen], np(W), [.].
sent(d, s(X, Y, Z, [aen], W)) --> np(X), mv(Y), co
mpar(Z), [aen], np(W), [.].
sent(d, s(X, [aer], Y, Z, [aen], W)) --> np(X), [a
er], mf(Y), compar(Z), [aen], np(W), [.].
sent(d, s(X, Y, [lika], [mycket], [som], Z)) --> n
p(X), mv(Y), [lika], [mycket], [som], np(Z), [.].
sent(d, s(X, [aer], [lika], Y, [som], Z)) --> np(X
), [aer], [lika], a(Y), [som], np(Z), [.].
sent(d, s(X, [aer], Y, [gg_saa], Z, [som], W)) -->
 np(X), [aer], num(Y), [gaanger], [saa], a(Z), [so
m], np(W), [.].
sent(d, s(X, [aer], Y, [gg], Z, [aen], W)) --> np(
X), [aer], num(Y), [gaanger], compar(Z), [aen], np
(W), [.].
sent(q, s(X, [aer], mf(hur), Y)) --> [hur], a(Y),
[aer], np(X), [?].
sent(q, s(X, [aer], mf(hur_maanga, Y), Z)) --> [hu
r_maanga], unit(Y), a(Z), [aer], np(X), [?].
sent(q, s(X, Y, mf(hur_maanga, Z))) --> [hur_maang
a], unit(Z), mv(Y), np(X), [?].
sent(q, s(Z, Y, X)) --> mf(X), mv(Y), np(Z), [?].
sent(q, s(X, Y, mf(hur_mycket))) --> [hur_mycket],
 mv(Y), np(X), [?].
sent(q, s(X, Y, mf(vad))) --> [vad], mv(Y), np(X),
 [?].
sent(q, s(X, [har], Y, mf(vilken))) --> [vilken],
mn(Y), [har], np(X), [?].
sent(q, s(X, Y, [aer], mf(vad))) --> [vad], [aer],
 nps(X), mn(Y), [?].
sent(q, s(X, Y, [aer], mf(vilken))) --> [vilken],
[aer], nps(X), mn(Y), [?].
sent(q, s(X, [aer], Y)) --> [aer], np(X), a(Y), [?
].
sent(q, s(X, Y, i(Z))) --> mv(Y), np(X), i(Z), [?]
.
sent(q, s(X, Y, Z)) --> mv(Y), np(X), mf(Z), [?].
sent(q, s(X, [aer], Y, Z)) --> [aer], np(X), mf(Y)
, a(Z), [?].
sent(q, s(X, [har], id(Y), Z)) --> [har], np(X), i
d(Y), mn(Z), [?].
sent(q, s(X, [aer], Z, [aen], W)) --> [aer], np(X)
, compar(Z), [aen], np(W), [?].
sent(q, s(X, Y, Z, [aen], W)) --> mv(Y), np(X), co
mpar(Z), [aen], np(W), [?].
```

```
/* lexikon */
np(X) --> [X], {isn(X)}.
isn(sven).
isn(lisa).
isn(stolen).
isn(bilen).
isn(stenen).
isn(mannen).
nps(X) --> [X], {isns(X)}.
isns(hans).
gen(hans, hans).
isns(Y) :- isn(X), gen(X, Y).
gen(X, V) :- name(X, Y), name(s, Z), append(Y, Z,
W), name(V, W).
mn(X) --> [X], {ismn(X)}.
ismn(bredd).
ismn(hoejd).
ismn(tjocklek).
ismn(storlek).
ismn(tyngd).
ismn(vidd).
ismn(laengd).
ismn(kvalitet).
ismn(styrka).
ismn(kvantitet).
ismn(vikt).
ismn(aalder).
ismn(kostnad).
ismn(volym).
ismn(temperatur).
a(X) --> [X], {isa(X)}.
isa(stark).
isa(djup).
isa(svag).
isa(klok).
isa(dum).
isa(tung).
isa(laett).
isa(billig).
isa(dyr).
isa(snabb).
isa(bred).
isa(smal).
isa(laang).
isa(kort).
isa(stor).
isa(liten).
isa(hoeg).
isa(laag).
isa(tjock).
isa(gammal).
isa(ung).
isa(varm).
isa(kall).
isa(bra).
isa(daalig).
i(X) --> [X], {isi(X)}.
isi(mycket).
isi(lite).
isi(foega).
id(X) --> [X], {isid(X)}.
isid(stor).
```

```
/* dimension-unit relations */
du(length, meter).
du(height, meter).
du(height, fot).
du(depth, meter).
du(temperature, grader).
du(cost, kronor).
du(weight, kilo).
du(age, aar).
du(volume, liter).
du(time, timmar).
du(time, minuter).
/* dimensional meaning of certain adjectives (tran
slations) */
b(hoeg, height).
b(gammal, age).
b(laang, length).
b(stor, size).
b(varm, temperature).
b(bred, width).
b(djup, depth).
b(tjock, thickness).
/* derivation of dimenional(unmarked) adjective */
b(X, D) :- b(X, D, more), isa(X).
/* dimenional meaning of certain verbs */
vb(kostar, cost).
vb(vaeger, weight).
vb(raeknar, size).
vb(rymmer, volume).
vb(maeter, length).
/* meaning of certain nouns */
b(kostnad, cost).
b(vikt, weight).
b(storlek, size).
b(laengd, length).
b(hoejd, height).
b(aalder, age).
/* derivation of form and meaning of comparatives
*/
cb(tyngre, weight, more).
cb(aeldre, age, more).
cb(yngre, age, less).
cb(laengre, length, more).
cb(stoerre, size, more).
cb(mindre, size, less).
cb(hoegre, height, more).
cb(mer, size, more).
cb(laegre, height, less).
cb(X, Y, Z) :- atocompar(W, X), b(W, Y, Z).
compar(X) --> [X], {iscompar(X)}.
atocompar(X, Y) :- isa(X), name(X, Z), name(are, W
), append(Z, W, V), name(Y, V).
iscompar(Y) :- isa(X), atocompar(X, Y).
iscompar(aeldre).
iscompar(tyngre).
iscompar(stoerre).
iscompar(laengre).
iscompar(laegre).
iscompar(hoegre).
```

```
/* translation into logic */
to_log(F, L) :- F=s(X, [aer], Y), L=l(X, D, E), b(
Y, D, E).
to_log(F, L) :- F=s(X, Y, i(Z)), L=l(X, D, E), vb(
Y, D), ib(Z, E).
to_log(F, L) :- F=s(X, [har], id(Y), Z), L=l(X, V,
 E), ib(Y, E), b(Z, V).
to_log(F, L) :- F=s(X, [har], Y, mf(Z, M)), L=l(X,
 V, Z, M), b(Y, V), du(V, M).
to_log(F, L) :- F=s(X, [aer], mf(Z, M), U), L=l(X,
 Y, Z, M), b(U, Y), du(Y, M).
to_log(F, L) :- F=s(G, Y, [aer], mf(Z, M)), L=l(X,
 V, Z, M), gen(X, G), b(Y, V), du(V, M).
to_log(F, L) :- F=s(X, Y, mf(Z, U)), L=l(X, D, Z,
U), vb(Y, D), du(D, U).
to_log(F, L) :- F=s(X, [aer], Z, [aen], Y), L=l(X,
 D, E, Y), cb(Z, D, E).
to_log(F, L) :- F=s(X, Y, I, [aen], Z), L=l(X, D,
E, Z), vb(Y, D), cb(I, D, E).
to_log(F, L) :- F=s(X, [aer], mf(Y, M), Z, [aen],
W), L=l(X, D, E, Y, M, W), cb(Z, D, E), du(D, M).
to_log(F, L) :- F=s(X, [aer], [lika], Y, [som], Z)
, L=l(X, D, E, like, Z), b(Y, D, E).
to_log(F, L) :- F=s(X, [aer], [lika], Y, [som], Z)
, L=l(Z, D, E, like, X), b(Y, D, E).
to_log(F, L) :- F=s(X, Y, [lika], [mycket], [som],
 Z), L=l(X, D, more, like, Z), vb(Y, D).
to_log(F, L) :- F=s(X, [aer], Y, [gg_saa], Z, [som
], W), L=l(X, D, E, Y, times, W), b(Z, D, E).
to_log(F, L) :- F=s(X, [aer], Y, [gg], Z, [aen], W
), L=l(X, D, E, Y, times, W), cb(Z, D, E).
/* implicational rule about old and age for humans
 */
exp(X, D, E) :- l(X, D, E), !.
exp(X, D, E) :- l(X, age, Y, _), Y>60, human(X), E
=more.
human(sven).
human(hans).
human(lisa).
/* inferences */
/* transitivity */

exp(X, D, E, Z) :- l(X, D, E, Y), l(Y, D, E, Z), p
rint([as, X, has, E, D, than, Y, and, Y, E, than,
Z]).
exp(X, D, E, Z) :- l(X, D, E, Y), exp(Y, D, E, Z),
 print([as, X, has, E, D, than, Y, and, Y, E, than
, Z]).
exp(X, D, E, Z) :- l(X, D, E, Y), exp(Y, D, E, lik
e, Z), print([as, X, has, E, D, than, Y, and, Y, i
s, like, Z]).
exp(X, D, E, Z) :- exp(X, D, E, like, Y), l(Y, D,
E, Z), print([as, X, is, like, Y, and, Y, has, E,
D, than, Z]).
exp(X, D, E, like, Y) :- l(X, D, E, like, Y), !.
exp(X, D, E, like, Z) :- l(X, D, E, like, Y), l(Y,
 D, E, like, Z), print([as, X, is, like, Y, and, Y
, is, like, Z]).
exp(X, D, E, like, Z) :- l(X, D, E, like, Y), exp(
Y, D, E, like, Z), print([as, X, is, like, Y, and,
```

```
/* interactive commands */
logic(X) :- sent(T, F, X, []), to_log(F, L), print
(L), nl.
synt(X) :- sent(T, F, X, []), print(F), nl.
accept(X) :- sent(T, F, X, []), to_log(F, L), asse
rt(L), print(jaha), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [aer],
 mf(hur_maanga, Z), W), l(Y, D, N, Z), b(W, D), pr
int([N, Z]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, mf(
hur_maanga, W)), l(Y, D, N, M), vb(Z, D), print([N
, M]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [har],
 Z, mf(vilken)), l(Y, V, N, U), b(Z, V), print([N,
 U]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, mf(
hur_mycket)), l(Y, D, N, U), vb(Z, D), print([N, U
]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, mf(
vad)), l(Y, D, N, U), vb(Z, D), print([N, U]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [aer],
 mf(hur), Z), l(Y, D, N, U), b(Z, D), print([N, U,
 Z]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, [ae
r], mf(vad)), l(G, D, N, U), b(Z, D), gen(G, Y), p
rint([N, U]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, [ae
r], mf(vilken)), l(G, D, N, U), b(Z, D), gen(G, Y)
, print([N, U]), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [aer],
 Z), exp(Y, V, E), print(ja), b(Z, V, E), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [aer],
 Z), exp(Y, V, P), P\=E, b(Z, V, E), print(nej), n
l.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, i(W
)), vb(Z, V), ib(W, E), exp(Y, V, E), print(ja), n
l.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, i(W
)), vb(Z, V), ib(W, more), exp(Y, V, less), print(
nej), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [har],
 id(Z), V), b(V, D), ib(Z, E), exp(Y, D, E), print
(ja), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, Z, mf(
V, W)), l(Y, D, V, W), vb(Z, D), print(ja), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [aer],
 mf(V, M), Z), l(Y, D, V, M), b(Z, D), print(ja),
nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [aer],
 Z, [aen], W), exp(Y, D, E, W), cb(Z, D, E), print
(ja), nl.
answer(X) :- sent(T, F, X, []), T=q, F=s(Y, [aer],
 [hur_mycket], Z, [aen], W), l(Y, D, E, N, M, W),
b(Z, D, E), print(N, M), nl.
```

A note on Kammu perception verbs

Jan-Olof Svantesson

In this note I will describe a semantic field, perception verbs in the Austroasiatic language Kammu. I will rely heavily on the format and concepts developed by Åke Viberg (1982a) and I have also used his questionnaire (Viberg 1982b) to elicit perception verbs in Kammu.

Viberg classifies basic perception verbs in two ways, according to the five sense modalities (sight, hearing, touch, taste, smell), and according to the "dynamic meaning" of the verb (activity, experience and copulative). Activity and experience are experiencer-based (have the experiencer as subject) while copulative is source-based. Activity refers to a consciously controlled process, while experience refers to a state that is not controlled. This is reflected in the fact that while activity verbs can be uttered in imperative or prohibitive form, experience verbs cannot. For instance, in Kammu the sentences sŋə́əy "Look!" or táa kmñèŋ "Don't listen!" are possible, but *kùuñ "See!" or *táa mèc "Don't hear!" are not.

The three dynamic meanings are exemplified for "sight" by the following sentences (Viberg 1982b):

A(ctivity):	Peter looked at the birds
E(xperience):	Peter saw the birds
C(opulative):	Peter looked happy

Thus, fifteen basic terms are obtained. They are illustrated by the following sentences, which were translated by a Kammu informant (Kàm Ràw) from Viberg's questionnaire:

Sight:	A	Làay sŋə́əy síim Làay look bird	"Làay looked at the bird."
	E	Làay kùuñ síim Làay see bird	"Làay saw the bird."
	C	Làay mɨ́an cə̀ə lə̀ hrñɨam Làay like PTC happy	"Làay looked happy."

Hearing: A Lòay kmñèŋ síim (yàam) / Lòay listen bird (call) — "Lòay listened to the bird (sing)."

E Lòay mèc síim (yàam) / Lòay perceive bird (call) — "Lòay heard the bird (sing)."

C Lòay mían càə lè hrñíam / Lòay like PTC happy — "Lòay sounded happy."

Touch: A Lòay míap téep tèe / Lòay feel shirt RFL — "Lòay felt his shirt."

E Lòay mèc klòaŋ yèt tòa kntrùum cìaŋ tèe
Lòay perceive stone stay at underside foot RFL
"Lòay felt a stone under his foot."

C téep lmèɛt lntrèen / shirt smooth EXP — "The shirt felt smooth."

Taste: A Lòay cìim sŋmàh / Lòay taste food — "Lòay tasted the food."

E Lòay mèc sŋmàh cáŋ máar / Lòay perceive food bitter salt — "Lòay tasted salt in the food."

C sŋmàh lòm / food tasty — "The food tasted good."

sŋmàh cáŋ máar / food bitter salt — "The food tasted of salt."

Smell: A Lòay hmʔɨr sŋmàh / Lòay smell food — "Lòay smelt the food."

E Lòay mèc ráaŋ sʔóɔŋ hʔɨ́r tòa prì
Lòay perceive flower tree smell in forest
"Lòay smelt flowers in the forest."

C sŋmàh hʔɨ́r / food smell-good — "The food smelt good."

sŋmàh hʔú / food stink — "The food stank."

sŋmàh hʔɨ́r hóɔm / food smell onion — "The food smelt of onions."

The perception verbs are summarized in the following table:

	activity	experience	copulative
sight	sŋə́əy	kùuñ	mɨ́an
hearing	kmñèŋ	mèc	mɨ́an
touch	mɨ́ap	mèc	-
taste	cìim	mèc	-
smell	hmʔɨ̀r	mèc	hʔɨ́r, hʔú

For activity, there are five different verbs, according to the sense involved, but for experience there are only two, one for "see" and one for the remaining four senses. The copulatives are more complicated: for "look" and "sound" the verb mɨ́an "(be) like" combined with the aspect particle cə̀ə is often used. This particle basically denotes that the speaker does not know whether the state or action denoted by the following verb takes place or not. (Thus it is often used for future tense, but this is not the case here; see also Svantesson 1984.) The meaning of the combination mɨ́an cə̀ə is thus something like "seem".

For touch and taste there are no specific copulative verbs, and various paraphrases are used. For smell there are two basic terms, one for "smell good" and one for "smell bad". These and other hyponyms will be described below.

Polysemy

The dynamic meanings experience and activity are lexicalized by two different verbs for each sense modality and in no case are these two verbs morphologically related to each other. This contrasts with Vietnamese, the only Austroasiatic language in Viberg's sample where the experience verbs are formed from activity verbs by the addition of a serial verb.

Viberg has shown that there is a polysemy hierarchy within the sense modalities in different languages:

sight > hearing > touch > smell / taste

This means that "a verb having a basic meaning belonging to a sense modality higher (to the left) in the hierarchy, can get an extended meaning that covers some (or all) of the sense modalities lower in the hierarchy" (Viberg 1982a:15).

As seen above, the activity verbs are lexicalized differently for each sense modality, but for experience, the verb mèc is used for all senses except sight. If the context allows a choice between sense modalities, mèc is usually interpreted as "hear", which is thus its basic meaning. Thus the Kammu data fit in with this hierarchy, as well as with Viberg's observation that polysemy is most common among the experiences. Languages with the same experience polysemy pattern as Kammu (i.e. with a separate verb for sight and with the verb for hearing extended to the remaining senses) are common in Viberg's language sample and include Hausa and Turkish.

The only perception verbs which are morphologically related in Kammu are hmʔɨr "smell (activity)" and hʔɨ́r "smell (copulative)". The infix -m- regularly forms causatives in Kammu (see Svantesson 1983), and the formal relation between copulative and activity verbs is the same as that between an intransitive verb and its causative; the first is constructed **patient/source - V**, and the second **agent - V - patient/source**.

Hyponymy

Hyponyms for "look" (activity) and for "smell" (copulative) seem to be especially common, and I have collected some of these here.

For "look", the following hyponyms have been noted:

cŋkléɛr	"look through a hole, look into, peep"
kléɛr	= cŋkléɛr
smdíŋ	"stare at"
smfìik	"glance at, spy on"
smlə̀ŋ	"look to the side"
smŋə̀h	"look upwards" (cf. ŋə̀h "bent upwards")
smʔyée	"look with one eye" (cf. ʔyée "one-eyed")

Many of these words have the minor syllable sm-, which in the last two cases seems to be a kind of derivational prefix.

For copulative "smell" there is no indigenous neutral verb, but one has to specify if the smell is good (hʔɨ́r) or bad (hʔú). The Lao loanword sáap seems to be neutral. There are several other words for different kinds of smell:

hʔɨ́r	"smell good"
hʔú	"smell bad"
ŋéɛl	"to stink (of animals, dirty clothes, etc.)"
pʔìal	"to stink (of fish)"
sáap	"smell" (Lao loan)
ʔwɨ́as	= ŋéɛl

These words function both as verbs and as nouns (which is not very common in Kammu):

ká pʔìal
fish stink "The fish stinks."

ò mèc pʔìal ká
I perceive stench fish" "I smell the stench of fish."

Bibliography

Svantesson, Jan-Olof. 1983. **Kammu phonology and morphology**. Lund: Gleerup.

Svantesson, Jan-Olof. 1984. "Sketch of the tense-mood-aspect system of Kammu". In **Selected working papers from the tense-mood-aspect project**, ed. Östen Dahl and Dora Kós-Dienes, Stockholm, p. 44-48.

Viberg, Åke. 1982a. **The verbs of perception. A typological study**. Stockholm: Institute of Linguistics

Viberg, Åke. 1982b. **Questionnaire: verbs of perception**. Stockholm: Institute of Linguistics.

Structures prosodiques du suédois et du français: le cas de la focalisation et du contraste

Paul Touati

Thèse à paraître

Le but principal des recherches menées dans ce travail de thèse est de décrire les structures prosodiques dans le cadre du problème linguistique posé par la focalisation et le contraste. Dans cette optique, nous avons:

1. Procéder à une description préalable de certains aspects de la prosodie du suédois et du français dans une perspective essentiellement phrastique.

2. Etudier les rapports entre les structures prosodiques, syntaxiques et énonciatives à travers les variations paramétriques de durée et de fo.

3. Analyser ces variations paramétriques dans le cadre du modèle mis au point à Lund.

1. A la manière de Gårding dans ses travaux [Gårding, 1981, 1984], nous avons opté pour une approche multilangagiaire. Notre choix s'est porté sur deux langues non-apparentées, le suédois et le français.

L'interêt de la comparaison du suédois et du français comme langues de réference n'échappera à personne. Les typologies lorsqu'elles se réfèrent aux langues européennes qui nous sont plus ou moins familières ne manquent pas de citer la prosodie du suédois et du français comme exemples illustratifs de systèmes prosodiques différents [Garde,1968; Hyman,1977].

Mais notre comparaison ne se veut pas être un simple répertoire des ressemblances et différences phonologiques entre les deux langues mais une analyse pluriparamétrique de la substance phonétique (durée et fo) basée sur les données d'un problème linguistique précis.

De plus choisir deux langues non-apparentées, c'est évité, nous l'espérons, l'écueil que Rossi met en évidence lorsqu'il montre combien les divergences d'opinion de Karcevskij, Danes et de Groot sur les rapports entre la syntaxe et l'énonciation étaient liées à la langue qui leur servait de base de reflexion [Rossi,1985:137].

2. La prosodie constitue un vaste champ de recherches. Notre étude se situe dans le cadre des rapports entre les structures prosodiques, syntaxiques et énonciatives [cf."Semantics, syntax and prosody", Utrecht,1983]. Des limites plus claires lui sont posées par la nature du problème étudié, c'est à dire la manière dont l'acte d'énonciation constitué par la rhématisation d'un élément de l'énoncé influence la structure prosodique et indirectement la liaison de congruence entre les structures citées-ci dessus.

Deux types de rhématisation, tenus comme différents dans notre hypothèse de départ, sont envisagés; il s'agit de la focalisation et du contraste où la focalisation a une fonction d'identification alors que le contraste a une fonction de selection exclusive [Rossi et al, 1981:230; Nølke, 1983:150-151]. De manière générale, nous avons essayé de décrire, avec l'aide de la méthode mise au point par Bruce [Bruce,1977: 21 sq], les propriétés prosodiques d'énoncés neutres, d'énoncés avec focalisation et d'énoncés avec contraste dérivés à partir *d'une même phrase* (d'une même structure syntaxique).

Les propriétés de ces énoncés sont saisies en termes de variations des paramètres acoustiques de durée et de fo. Une fois normalisées, les variations de durée constituent ce que nous avons qualifié de profils temporels [cf. Touati, 1984]. Les variations locales de fo dues aux accents sont quantifiés. Les variations globales de fo sont exploitées qualitativement.

3. L'analyse des données acoustiques, surtout celle de fo, est effectué à l'aide des unités (points-clefs, grille intonative, pivots...) proposées dans le modèle mis au point à Lund [Gårding,1984 ; Gårding et al,1982, servant de point de départ à notre analyse]. Ces unités fonctionnent de la manière suivante.

Chaque accent a une représentation phonologique intermédiaire sous forme de points-clés Hauts ou Bas. La combinaison de ces points contribue à doter les accents de contours montants ou descendants. Aux points-clés correspondent des maxima et des minima de fo. Les contours se traduisent donc par des chutes et des montées de fo dont l'ampleur est estimée à l'aide du rapport maximum de fo sur minimum de fo.

Les lignes de base et de sommet qui relient respectivement les minima et les maxima des points-clés majeurs composent une grille dont la direction est censée fournir une claire indication quant au mouvement globale de l'intonation. Les lignes ont aussi pour vertu de permettre la comparaison entre les différentes parties de la grille -les différents groupes intonatifs- en ce qui concerne leur expansion fréquentielle à l'aide des traits comme par exemple [± ample] ou [±comprimée]. Les pivots sont des lieux de rupture de la grille. Une étroite liaison entre ces unités et les fonctions prosodiques est postulée.

Ce travail de thèse n'étant pas totalement achevé, les résultats présentés dans le cadre de ce résumé sont des généralisations préliminaires.

Mais de manière générale, on peut affirmer que:

1. Il existe des différences phonologiques entre les deux langues qu'il serait trop long d'exposer ici.

2. Outre ces différences, on remarque qu'une similitude phonologique dans les deux langues (accentuation lexicale finale) se traduit par des différences fondamentales dans les réalisations paramétriques.

3. La réalisation paramétrique des accents est sensible à leur catégorie (accent de mot/de groupe, accent focal/contraste) et se réalise différement dans les deux langues.

4. Les grilles intonatives chargées de saisir les variations acoustiques globales ont en commun un certain nombre de traits dont la distribution varie cependant selon la langue, la position de l'accent focal ou de contraste et le type d'énoncé.

4. Une autre différence fondamentale entre les deux langues réside dans le fait que les paramètres de durées et de fo entretiennent une relation de dépendance relativement forte en suédois alors qu'en français, ils seraient plutôt indépendants l'un de l'autre.

En ce qui concerne les effets prosodiques de la focalisation et le contraste, on observe que:

1. Dans le deux langues, la rhématisation aboutit au fait que la congruence entre structure syntaxique et la structure prosodique n'est pas maintenue. Elle disparait au profit de celle entre structure énonciative et prosodique.

2. Les vecteurs prosodiques de la focalisation et du contraste sont les accents dit "accent focal" et "accent de contraste". L'existence de ces catégories comme catégories différentes n'est pas clairement attestée. En français, il est ainsi difficile de dériver à partir des réalisations paramétriques une catégorie "accent focal". Les manisfestations de la catégorie "accent de contraste" sont en revanche claires. En suédois, les catégories "accent focal" et "accent de contraste" sont acoustiquement réalisées de la même manière. Elles ne diffèrent pas qualitativement mais quantitativement.

3. La réalisation paramétrique des accents de mot (pour le suédois) ou de groupe (pour le français) est sensible à leur position par rapport à l'accent focal/de contraste (position postfocale et position préfocale).

Références

Bruce, G. 1977. Swedish word accent in sentence perspective. Travaux de l'institut de Lund XII. Lund: Gleerups.

Garde, P. 1968. L'accent. Presses Universitaires de France, Paris.

Gårding, E. 1981. Contrastive prosody: A model and its application. Special lecture to the AILA congress 1981. Proceedings II, eds. Sigurd and Svartvik, Studia Linguistica 35, 1-2, 81, Gleerups: 146-166.

Gårding, E., Botinis, A. and Touati, Paul. 1982. A comparative study of Swedish, Greek and French Intonation. Working Papers 22, Departement of Linguistics, Lund University: 137-152.

Gårding, E. 1984. Comparing intonation. Working Papers 27, Departement of Linguistics, Lund University: 75-96.

Hyman, L. 1977. Studies in stress and accent. University of Southern California, Los Angeles.

Nølke, H. 1983. Remarques sur la focalisation. Revue Romane, numéro special 24, 147-165.

Rossi, M. 1985. L'intonation et l'organisation de l'énoncé. Phonetica 42:135-153.

Rossi, M. Di Cristo, A. Hirst, D. Martin, Ph. et Nishimuda, Y. 1981. L'intonation de l'acoustique à la sémantique. Paris: Klincksieck.

Touati, P. 1984. Profils temporels en suédois et en français. Dans Elert, C-C. Johansson, I. and Strangert, E. (Eds) Nordic Prosody III, Acta Universitatis Umenis 59, University of Umeå, 230-242.

CURRENT RESEARCH PROJECTS

Commentator, a computer system simulating verbal behavior

Bengt Sigurd

Computer simulation of dialogue and communication

Bengt Sigurd

Manuals and instructions with particular reference to the computer environment
(Swedish Council for Research in the Humanities and Social Sciences)

Bengt Sigurd, Boel de Geer

Language and brain
(Swedish Council for Research in the Humanities and Social Sciences)

Eva Gårding, Christina Dravins

Central features in Scandinavian syntax
(Nordic Committee for Coordinating Research in the Humanities)

Elisabet Engdahl

Phonological and phonetic studies of vowel articulation in Bulgarian

Thore Pettersson, Sidney Wood

Language disordered children's reading and spelling
(Swedish council for Research in the Humanities and Social Sciences)

Kerstin Nauclér, Eva Magnusson

Reading and spelling difficulties. A search for causes.
(Swedish Council for Research in the Humanities and Social Sciences)

Kerstin Nauclér, Eva Magnusson

Diphtongization versus vowel duration in Southern Swedish

Kurt Johansson

Prosodic parsing for Swedish speech recognition
(National Swedish Board for Technical Development)

Gösta Bruce, David House (Lund)
Björn Lindblom, Francisco Lacerda (Stockholm)

From text to prosody
(The Bank of Sweden Tercentenary Foundation)

Robert Bannert

LECTURES AND PUBLICATIONS ELSEWHERE

1985

Bannert, R	Fokus, Kontrast und Phrasenintonation im Deutschen. Zeitschrift für Dialektologie und Linguistik 52/3
Bannert, R.	Towards a model for German prosody. Folia Linguistica XIX, 321-341
Engdahl, E.	Parasitic gaps, resumptive pronouns and subject extractions. Linguistics 23:1
Gårding, E.	An intonation model and its anchors. Institut de phonétique, Aix-en-Provence, July
Gårding, E.	A phrase-based model of intonation. Institut für Phonetik und sprachliche Kommunikation, München, July
Gårding, E.	& House, D. Frasintonation, särskilt i svenska. In Allén et al. (eds.): Svenskans beskrivning 15, Förhandlingar vid sammankomst för att dryfta frågor rörande svenskans beskrivning, 205-221. Göteborgs universitet
Gårding, E.	En intonationsmodell och dess förankringar. Tal, ljud, hörsel 2, GULING 13, Göteborgs universitet
Horne, M.	English sentence stress, grammatical functions and contextual coreference. Studia Linguistica 39, 51-66
House, D.	Betydelsen av spektrala förändringar för kategorisering av tonala mönster. Abstract in Ahlsén, Allwood & Hjelmquist (eds.): Tal, ljud, hörsel 2, p 168. Göteborgs universitet
House, D.	Kompensatoriska perceptionsstrategier bland hörselskadade lyssnare. Abstract in Ahlsén, Allwood & Hjelmquist (eds.): Tal, ljud, hörsel 2, p 168. Göteborgs universitet
House, D.	Compensatory use of acoustic speech cues and language structure by hearing-impaired listeners. Paper presented at the conference "Communication and handicap - aspects of psychological compensation and technical aids", Stockholm, 3-4 June
House, D	Perception of fundamental frequency movement in speech. Poster presented at the Third international conference on event perception and action, Uppsala University, 24-29 June
House, D.	Implications of spectral changes for categorising tonal patterns in speech. Paper presented at the British Society of Audiology short papers meeting on experimental studies of hearing and deafness. University of Hull, 8-9 July

House, D. Speech perception, intonation and hearing impairment. Paper presented at the "Tal-Ljud-Hörsel" summer school, Särö, 17-24 August

Pettersson, T. & Persson, G. Tillåt formerna "dom" och "mej". Dagens Nyheter, 15 March

Pettersson, T. Dags att skriva "dom". Paper presented at Språkvårdsfestival, City Library, Lund, 20 November

Pettersson, T. & Wood, S. Spektrografičeskie i model'nye issledovanija bolgarskix glasnyx (Spectrographic and model studies of the Bulgarian vowels). Paper presented at III Bălgaro-skandinavski simpozium po bălgaristika, Varna, Bulgaria. September

1986

Bannert, R. Acoustic-phonetic investigations of the intonation of German and Danish: similarities and differences. ARIPUC 20, 45-66

Engdahl, E. Constituent questions: the syntax and semantics of questions with special reference to Swedish. Reidel, Dordrecht

Gårding, E. The structure and manifestation of intonation & Learning and teaching problems in the light of a model for prosody. Papers presented at Nankai University, Tjanjin, and Beijing University, Beijing, May

Gårding, E. Intonation across several languages. Lecture series, Institute of Acoustics, Academia Sinica, Beijing. May-June

Gårding, E. Superposition as an invariant feature of intonation. In Perkell, J.& Klatt, D.H.(eds.): Invariance and variability in speech. L.Erlbaum Assoc.Publ., Hillsvale, N.J.

Horne, M. Information focus: assignment and phonological implications. In Evensen, L. (ed.): Nordic research in text linguistics and discourse analysis, 155-69. Tapir, Trondheim

House, D. Compensatory use of acoustic speech cues and language structure by hearing-impaired listeners. Forthcoming in Nilsson, L.G.& Hjelmquist, E.(eds.): Communication and handicap - aspects of psychological compensation and technical aids. North Holland, Amsterdam

House, D. & Gårding, E. Phrasing in some scandinavian dialects. Paper presented at Nordic prosody IV, Middelfart, 6-8 June

Pettersson, T. Dags att skriva "dom". Sydsvenska Dagbladet, 11 January

Pettersson, T. Stavning av pronomina. Sydsvenska Dagbladet, 27 January

Pettersson, T. Nekotorye nabljudenija nad voprosom o grammatičeskoj interferencii (na materiale švedskogo i russkogo jazykov). Paper presented at the Inter-Scandinavian conference "Språklig interferens (nordiska språk vis-à-vis ryska), Helgenæs, Denmark, 25-27 April

Pettersson, T. Kasus i ryska och andra språk. Århus University, 28 April

Pettersson, T. & Bílý, M. Ergativity in Dyirbal. In Mey, J.L. (ed): Language and discourse: test and protest. A festschrift for Peter Sgall (=Linguistic and literary studies in Eastern Europe 19), 143-161. John Benjamins Publishing Company, Amsterdam

Söderbergh, R. From communication as language to language as communication. Keynote address in Precursors of early speech, eds. Lindblom, B. & Zetterström, R. (Wenner-Gren Center International Symposium Series, v. 44). Stockton Press, Southampton

Wood, S. & Pettersson, T. Spektrografičeskie i model'nye issledovanija bolgarskix glasnyx (Spectrographic and model studies of the Bulgarian vowels). Invited lecture, Second International Congress on Bulgarian Studies, Sofia. Proceedings in press. May

Wood, S. & Pettersson, T. Vowel reduction in Bulgarian and its implications for theories of vowel production: a review of the problem. Folia Linguistica (in press)

Wood, S. The acoustical consequences of lip, tongue and larynx articulation i palatal vowels. In Journal of the Acoustical Society of America (in press)

Wood, S. The role of the pharynx in speech, Chapter in Otolaryngology - head and neck surgery (vol. 2 The larynx and pharynx), ed. Painter, C., C.V. Mosby Co., St. Louis

RECENT NUMBERS AND CONTENTS

WP 28 1985	EVA GÅRDING, PAUL KRATOCHVIL, JAN-OLOF SVANTESSON & JIALU ZHANG	Tone 3 and Tone 4 discrimination in Modern Standard Chinese
	DAVID HOUSE	Implications of rapid spectral changes on the categorization of tonal patterns in speech perception
	DAVID HOUSE	Sentence prosody and syntax in speech perception
	DIETER HUBER	Swedish intonation contours in text-to-speech syntesis
	KERSTIN NAUCLÉR & EVA MAGNUSSON	Language disordered children's reading and spelling: preliminary results
	KJELL NORLIN	Pharyngalization in Cairo Arabic
	JANINA PETECKA	A study of question intonation in Polish

TRAVAUX DE L'INSTITUT DE LINGUISTIQUE DE LUND
FONDÉS PAR BERTIL MALMBERG
PUBLIÉS PAR EVA GÅRDING ET BENGT SIGURD

I	*Carl-Gustav Söderberg.* A Typological Study on the Phonetic Structure of English Words with an Instrumental-Phonetic Excursus on English Stress. 1959.
II	*Peter S. Green.* Consonant-Vowel Transitions. A Spectrographic Study. 1959.
III	*Kerstin Hadding-Koch.* Acoustico-Phonetic Studies in the Intonation of Southern Swedish. 1969.
IV	*Börje Segerbäck.* La réalisation d'une opposition de tonèmes dans des dissyllables chuchotés. Étude de phonétique expérimentale. 1966.
V	*Velta Rūķe-Draviņa.* Mehrsprachigkeit im Vorschulalter. 1967.
VI	*Eva Gårding.* Internal Juncture in Swedish. 1967.
VII	*Folke Strenger.* Les voyelles nasales françaises. 1969.
VIII	*Edward Carney.* Hiss Transitions and their Perception. 1970.
IX	*Faith Ann Johansson.* Immigrant Swedish Phonology. 1973.
X	*Robert Bannert.* Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage. 1976.
XI	*Eva Gårding.* The Scandinavian Word Accents. 1977.
XII	*Gösta Bruce.* Swedish Word Accents in Sentence Perspective. 1977.
XIII	*Eva Gårding, Gösta Bruce, Robert Bannert* (Eds). Nordic Prosody. 1978.
XIV	*Ewa Söderpalm.* Speech Errors in Normal and Pathological Speech. 1979.
XV	*Kerstin Nauclér.* Perspectives on Misspellings. 1980.
XVI	*Per Lindblad.* Svenskans sje- och tje-ljud (Some Swedish Sibilants). 1980.
XVII	*Eva Magnusson.* The Phonology of Language Disordered Children. 1983.
XVIII	*Jan-Olof Svantesson.* Kammu Phonology and Morphology. 1983.
XIX	*Ulrika Nettelbladt.* Developmental Studies of Dysphonology in children. 1983.

Numbers V, VII-XII, XIV-XIX are available from the publisher, Liber International, Room 405, S-205 10 MALMÖ, Sweden.

Numbers III, VI, XIII are available from the Department of Linguistics and Phonetics, Helgonabacken 12, S-223 62 LUND.

STUDIA LINGUISTICA

Eds Bengt Sigurd and Jan Svartvik

One volume of Studia Linguistica is published each year in two parts. The subscription price is 60 SEK (approximately $ 13). Orders may be sent to any international bookseller or to the publishers, CWK Gleerup, Box 1205, S-221 05 LUND, Sweden

PRAKTISK LINGVISTIK

Nr 1 - 1979	Eva Gårding & Robert Bannert Optimering av svenskt uttal (incl. cassette recording) out of print
Nr 2 - 1979	Kenneth Hyltenstam (Ed) Invandrarundervisning - aspekter på alfabetisering och den första nivån out of print
Nr 3 - 1979	Robert Bannert Ordprosodi i invandrarundervisningen 40:- (incl. cassette recording)
Nr 4 - 1979	Eva Wigforss Svensk pluralbildning hos finska invandrar-barn - en psykolingvistisk utvecklingsstudie 30:-
Nr 5 - 1981	Robert Bannert Svårigheter med svenskt uttal: inventering och prioritering 40:- (incl. cassette recording)
Nr 6 - 1981	Gösta Bruce (Ed) Tema prosodi - seminarieuppsatser i fonetik 30:-
Nr 7 - 1982	Don Kulick, Karina Vamling, Lars-Åke Henningsson Om tolkning 30:-
Nr 8 - 1983	Jan Fornell, Bengt Sigurd Commentator - ett mikrodatorbaserat forsknings-program för lingvister 30:-
Nr 9 - 1984	Thore Pettersson (Ed) Färgterminologi - seminarieuppsatser i allmän språkvetenskap 40:-
Nr 10 - 1985	Kerstin Nauclér, Eva Magnusson Språket i skrift - om normal och försenad utveckling av tal och skrift 60:-
Nr 11 - 1985	Bengt Sigurd, Jean-Jacques Berthout, Robin Cooper Språk på dator - Hur datorn kan konversera, förstå tolka morse och undervisa i franska 45:-

(Prices in Swedish Crowns)
Swedish Post Office Giro account 439 34 92 - 6

LINGUISTICS DEPARTMENT
STAFF

Chairman	Kurt Johansson
Director of studies	Thore Pettersson
Library	Boel de Geer
Office	Janna Geggus, Eva Körner, Britt Nordbeck

General Linguistics

Berthout, Jean-Jacques	BA; extra teacher
Cooper, Robin	PhD; Docent; Lecturer
Dooley, Sheila	BA; extra teacher
Engdahl, Elisabet	PhD; Research Fellow
de Geer, Boel	Fil Kand; Research Assistant
Henningsson, Lars-Åke	Fil Kand; extra teacher
Horne, Merle	MA; extra teacher
Malmberg, Bertil	Fil Dr; Professor Emeritus
Nordbeck, Britt	Secretary
Pettersson, Thore	Fil Dr; Docent; Lecturer
Sigurd, Bengt	Fil Dr; Professor; Head of Division; Dean
Karina Vamling	Fil Kand; Assistant, extra teacher
Research Students:	Jean-Jacques Berthout, Hans Dahlbäck, Sheila Dooley, Malgorzata Faust, Jan Fornell, Boel de Geer, Janna Geggus, Lars-Åke Henningsson, Merle Horne, Gisela Håkansson, Magnus Olsson, Emilio Rivano-Fischer, Karina Vamling

Phonetics

Bannert, Robert	Fil Dr; Docent; Research Assistant
Bruce, Gösta	Fil Dr; Professor; Head of Division
Dravins, Christina	Logoped; Research Assistant
Eeg-Olofsson, Mats	Fil Kand; Programmer
Geggus, Janna	Fil Lic; Secretary
Gårding, Eva	Fil Dr; Professor Emeritus
Hadding, Kerstin	Fil Dr; Professor Emeritus
House, David	BA; Assistant, extra teacher, Research Assistant

Johansson, Kurt	Fil Lic; Lecturer
Jönsson, Karl-Gustav	Technician
Körner, Eva	Secretary
Magnusson, Eva	Fil Dr; Logoped; Research Assistant
Nauclér, Kerstin	Fil Dr; Logoped; Research Assistant
Nettelbladt, Ulrika	Fil Dr; Logoped; extra teacher
Norlin, Kjell	Fil Kand; Research Assistant
Svantesson, Jan-Olof	Fil Dr; Research Assistant
Weimer, Kjell	Fil Kand; extra teacher
Willstedt, Ursula	Logoped; extra teacher
Wood, Sidney	Fil Dr; Research Fellow
Research Students:	Antonis Botinis, Anne-Christine Bredvad-Jensen, Christina Dravins, David House, Dieter Huber, Boris Larnert, Lili-Ann Rudberg, Gabriella Stenberg-Koch, Paul Touati, Yasuko Madsen, Kjell Weimer, Ursula Willstedt, Shi Bo

Voice and Speech Training

Norman, Lennart	Fil Kand; Voice and Speech teacher

Child Language Research

Bredvad-Jensen, Anne-Christine	Fil Kand; Research Assistant
Mårtensson, Bodil	Fil Kand; Secretary
Söderbergh, Ragnhild	Fil Dr; Professor; Head of Division

Address: Department of Linguistics and Phonetics
Lund University
Helgonabacken 12
S-223 62 LUND